PROTEOLYTIC NETWORK DYNAMICS IN BREAST CANCER AND TUMOR ASSOCIATED MACROPHAGES

A Dissertation
Presented to
The Academic Faculty

By

William Andrew Shockey

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Biomedical Engineering

Georgia Institute of Technology and Emory University

May 2019

PROTEOLYTIC NETWORK DYNAMICS IN BREAST CANCER AND TUMOR ASSOCIATED MACROPHAGES

Approved by:

Dr. Manu O. Platt, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Shelly Peyton
School of Chemical Engineering
University of Massachusetts Amherst

Dr. Melissa L. Kemp
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Eberhard O. Voit
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Christine Payne
School of Mechanical Engineering and Materials Science
Duke University

Date Approved: March 15, 2019

For the women in my life who have battled this terrible disease.

ACKNOWLEDGEMENTS

First, I would like to thank the members of my thesis committee for their valuable advice and guidance that made this thesis possible. Dr. Manu Platt, Dr. Melissa Kemp, Dr. Christine Payne, Dr. Shelly Peyton and Dr. Eberhard Voit all provided expertise to different aspects of this project and helped keep me focused on answering important questions.

For his tireless support over the last six years, I would like to thank my advisor Dr. Manu Platt. He has been deeply invested in my development as a person and scientist from the very beginning. I have never met someone more dedicated to their students and their personal, professional and academic development. Writing this dissertation required reviewing some of the first embarrassing experiments and presentations I did in the Platt Lab and it highlights how far I have come under Manu's mentorship. I will always be thankful to him for taking a chance on me.

I also owe Manu a debt of gratitude for assembling the group of extraordinary scientists and engineers of the Platt Lab, with whom I have been honored to work, play, eat, sing, dance, laugh, cry and celebrate during my time at Georgia Tech. Special thanks go out to my big brothers and sisters in the lab, Dr. Catera Wilder, Dr. Phil Keegan, Dr. Ivana Parker, Dr. Keon Young Park, Dr. Meghan Ferrall-Fairbanks and Dr. Chris Rivera, for helping me get started in the lab and welcoming me into the family that would carry me through my graduate career. Special thanks to my lab-twin, Dr. Akia Parks, who is a constant source of wisdom and entertainment and one of my most tight-knit friends. I have also been fortunate to welcome several younger graduate students into the lab. Simone

Douglas continuously inspires me with her drive to both achieve great science and help those less fortunate than themselves. Victor Omojola has already made great progress on a difficult project and I know he will carry on my legacy as the Louisiana boy that keeps the lab well fed. Darwin Carbajal is the newest member of the lab, but he has already impressed me with his focus and determination to become a great researcher and mentor.

In addition to these graduate students, I have been lucky enough to work several extremely talented and dedicated research scientists including Dr. Kristi Porter, Dr. Hannah Song, Dr. Chris Kieslich and Dr. Adeola Michael who have given me guidance on everything from dating and recipes to Western blotting and high-speed computing. I was also lucky to mentor a number of extraordinary high school and undergraduate students during my time at Georgia Tech, many of whom contributed work to this dissertation. I would like to thank Kendreze Holland, Ken Brandon, Alex Stewart, Aaron Thomas, Catelyn Evans, Eva Gatune, Dan Xu and Tiffanie Leeman for their countless hours of work.

Thank you to my undergraduate research mentor Dr. Todd Monroe for allowing me into his lab and inspiring to pursue research and graduate school. I would like to thank my high school teachers Dr. Kathy Morden and Dr. Peter Hawking for teaching the hardest physics and biology classes of my academic career—seriously—and showing me just how wonderful and crazy the world as we understand it can be. I would also like to thank my friends, both at Georgia Tech and around the world, who helped me get this far.

I would also like to thank my funding sources for making my training and research possible, including the National Institutes of Health, the Cell Manufacturing Technologies National Science Foundation Engineering Research Center, Project ENGAGES, and the Giglio family.

As I reach the end of my formal education, I have been reflecting on what the most important people in my life have taught me that allowed me to reach this milestone. My entire family showed me the value of unwavering support, always encouraging me to dream bigger and reach farther. My mom, Cindy, taught me the importance of emotional support and the ability to be independent and self-reliant. My dad, Bill, showed me the value of education in achieving my dreams. My sister, Jessie, inspired my curiosity and love of the natural world, which has stayed with me from playing with dinosaurs to playing with proteases. My sister, Danielle, showed me how far hard work and determination can take me. My brother-in-law Paul David taught me patience and quiet strength. My nephew, Jonah, and niece, Maggie, have brought me full circle, their curiosity reminding me how to see the natural world with child-like wonder. I will do my best to make sure they keep it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

BCA	Bicinchoninic acid protein assay
CAF	Cancer associated fibroblast
CatK	Cathepsin K
CatL	Cathepsin L
CatS	Cathepsin S
CatV	Cathepsin V
CCL2/MCP-1	Chemokine (C-C motif) ligand 2/Monocyte chemoattractant protein 1
CCL3/MIP-1α	Chemokine (C-C motif) ligand 3/Macrophage inflammatory protein 1-alpha
CHX	Cycloheximide
COPASI	Complex pathway simulator
CysC	Cystatin C
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
E64	trans-Epoxycuccinyl-L-leucylamido(4-guanidino)butane
ECM	Extracellular matrix
EGF	Epidermal growth factor
ER	Endoplasmic reticulum
FBS	Fetal Bovine Serum
FGF basic	Basic fibroblast growth factor
G-CSF	Granulocyte-colony stimulating factor
GAG	Glycosaminoglycans
HEK293T	Human Embryonic Kidney cell line
His-Tag	6x-histidine residues
IL-6	Interleukin 6
IL-8/CXCL8	Interleukin 8/Chemokine (C-X-C motif) ligand 8
M6P	Mannose 6-phosphate
MCF-10A	Michigan cancer foundation breast cancer cell line
MCF-7	Michigan cancer foundation breast cancer cell line
MCSF	Macrophage colony stimulating factor
MDA-MB-231	Monroe Dunway Anderson metastatic breast cell line
MM	Michaelis-Menten
MMP	Matrix metalloproteinase
mRNA	Messenger ribonucleic acid
MSC	Mesenchymal Stem Cell
ODE	Ordinary differential equations
Opti-MEM	Modified minimum essential medium
PACMANS	Protease-Ase Cleavage from MEROPS ANalyzed Specificities
PMA	Phorbol myristate acetate
PMA	phorbol 12-myristate 13-acetate
ProK	Procathepsin K
ProL	Procathepsin L
ProS	Procathepsin S

ProV	Procathepsin V
RAW264.7	Ralph and Williams murine macrophage cell line
Redox	Reduction-Oxidation
RPMI 1640	Monocyte media developed at Roswell Park Memorial Institute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAM	Tumor associated macrophage
Thp-1	Human monocyte cell line derived from pediatric patient with monocytic leukemia
TNFα	Tumor necrosis factor alpha
VEGF	Vascular endothelial growth factor
wt	wildtype

SUMMARY

Breast cancer metastasis is a complex process, promoted by a variety of cell types including cancerous mammary epithelial cells and stromal cells including fibroblasts, T cells and tumor associated macrophages (TAMs), which can make up 50% of tumor mass in late-stage, primary breast tumors. TAMs promote tumor growth, angiogenesis and metastasis through the release of inflammatory chemokines, growth factors and proteolytic enzymes including the cysteine cathepsins B, L, K, S and V. Macrophage secretion of these proteases and the endogenous cathepsin inhibitor, cystatin C, differs significantly between individuals, contributing to variability in patient outcomes and response to treatment. Cathepsins are potent catalysts of extracellular matrix (ECM) degradation, but have also been identified as regulators of cellular process that operate through the activation of signaling pathways and the cleavage of transmembrane receptor proteins, chemokines, and cellular adhesion molecules.

Cysteine cathepsins are components of a complicated regulatory proteolytic network, which includes multiple biochemical and biomechanical stimuli, endogenous protease inhibitors, and other proteases capable of activating or degrading other proteases while degrading substrates and inhibitors simultaneously. Cysteine cathepsins which include the most powerful mammalian collagenase and elastase, have become attractive pharmaceutical targets, motivating billions of dollars in research by pharmaceutical companies for a number of tissue destructive diseases. Despite the development of highly specific, tight binding inhibitors that have shown clinical efficacy in treating the target disease, no cathepsin inhibitor has yet received FDA approval due to the emergence of unexpected side effects in trial participants.

While cathepsins were once thought to be restricted to bulk protein turnover in lysosomes, researchers have begun to uncover the diverse tissue distribution, function and regulation of cathepsins both in healthy systems and in an ever-growing array of pathologies. Unraveling the complicated proteolytic web of interactions between enzymes, substrates and inhibitors will be critical to the development of cathepsin inhibitor treatments capable of suppressing cancer metastasis and other related diseases. Studying these mechanisms exclusively with experimental methods has been difficult, due to technological limitations, rapid reactions kinetics, promiscuity of proteases with substrates and the undetermined number of interactions and associated variables in these regulatory networks. This work combines experimental and computational methods to explore the effects of cathepsin inhibitors on intracellular and extracellular cathepsin dynamics in breast cancer.

Treating breast cancer cells with a cathepsin inhibitor has been shown to actually *increase* active cathepsins in those cells. We have since found this unexpected inhibitor response in macrophages and primary human tissue, which displays a great deal of person-to-person variability in endogenous active proteases and response to protease inhibitor treatment. This dissertation details the development of a mechanistic mathematical model, which revealed that this differential response to inhibitor treatment is driven by differences in intracellular environment and intrinsic properties of the enzymes. This model was also adapted to explore hypotheses about intracellular trafficking of cathepsin inhibitors and the impacts different classes of inhibitors can have on the same cellular cathepsin system that may be contributing causes to the unexplained side effects plaguing the pharmaceutical industry.

In addition to degrading the extracellular matrix and other protein substrates, some members of the cathepsin family have been shown to degrade other cathepsins, even in the presence of other substrates, a phenomenon known as "cathepsin cannibalism". This phenomenon was discovered using recombinant enzymes *in vitro*. In this work we have shown that cathepsin cannibalism can occur in cellular systems and produced a series of dynamic mathematical models to predict how these cathepsin-on-cathepsin interactions affect cathepsin levels and overall proteolysis in living cells. Cathepsin cannibalism could have a significant effect on systems where multiple cathepsins coexist, particularly in the tumor microenvironment where multiple cathepsin family members contribute to extracellular matrix degradation, facilitating tumor metastasis. Effective inhibition of matrix proteolysis in these systems will require an understanding of cathepsin-on-cathepsin interactions, since the addition of a specific cathepsin inhibitor could result in unexpected side effects due to the disruption of cannibalistic regulation.

This dissertation details the challenges inherent to studying systems of interacting biological agents in complex cellular systems and takes a hybrid experimental and mathematical approach to exploring these interactions. The data and models generated by this work will be useful in the creation of new experimental and mathematical tools to study these complex regulatory networks and inform the development of new cathepsin inhibitor treatments capable of helping people affected by a host of devastating diseases.

1 INTRODUCTION

1.1 Research Objectives and Specific Aims

The overall objective of this work was to develop experimental and mathematical tools capable of measuring and predicting proteolytic network dynamics in multicellular systems. The central hypothesis was that cathepsin inhibitors provoke unexpected side effects in cellular systems by disrupting cathepsin interactions with endogenous inhibitors, substrates and proteases.

Aim 1: Develop and characterize experimental tools to probe regulatory interactions between cathepsins, substrates and inhibitors in breast cancer and macrophages.

The hypothesis of this aim is that cathepsin profiles differ by cell type, resulting in cell-to-cell and person-to-person variation in proteolytic regulatory networks. Cathepsins are challenging enzymes to study experimentally as they are tightly regulated due to their potency, requiring consideration of pro and mature forms, activation, oxidation, substrates, loss of activity and denaturation. Work that contributes to this aim includes the characterization of the proteolytic profiles of macrophages (Chapter 3), the validation of mRNA for cathepsin and inhibitor overexpression (Chapter 5) and the purification of recombinant cathepsins (Chapter 6).

Aim 2: Predict multicellular active enzyme levels and generate hypotheses to identify unknown proteolytic network interactions using computational models.

The hypothesis of this aim is that one cathepsin species can affect the activity and expression of another cathepsin species through activation, degradation and competition

for substrates and inhibitors. Computational modeling is a useful technique for studying the interactive networks of proteases, particularly the protease states that can be difficult to resolve experimentally including complexes of cathepsins with inhibitors, substrates or other proteases; denatured or inactive proteases; amounts of degraded enzymes and substrates as they are removed from the system. Work that contributes to this aim includes the creation of a multicellular model of cathepsin S and K interactions in engineered cells (Chapter 4) and the development of a dynamic model of inhibitor induced increases in active cathepsins in breast cancer cells (Chapter 7).

1.2 Significance of Results

This work is innovative because it applies a hybrid computational and experimental approach to study a complex regulatory network that contributes to unexpected side effects of cathepsin inhibitor treatment in patients suffering from tissue destructive diseases that are cathepsin-mediated. Further, this work applies mechanisms previously documented *in vitro* and confirms their existence in cellular systems and generates tools enabling predictions of cathepsin dynamics in cellular systems. The long-term goal of this project is to define unknown regulatory functions of cathepsins, that when perturbed by inhibitors, provoke unexpected side effects that reduce their chances of being used clinically. Understanding these proteolytic network interactions will inform the design of new treatment modalities for people suffering from cancer and other cathepsin-mediated pathologies.

2 BACKGROUND

2.1 Breast cancer and tumor associated macrophages

2.1.1 Cancer metastasis and the tumor microenvironment

Breast cancer is the most common cancer among women and the second most deadly cancer in women [1]. Metastatic spread of secondary tumors to vital organs such as the lungs, liver and bone marrow is responsible for most deaths due to breast cancer [2, 3]. Cancer growth and metastasis are complex processes promoted by a variety of stromal cells, including tumor associated macrophages (TAMs) and cancer associated fibroblasts (CAFs) [4]. Breast cancer is a highly heterogeneous disease, manifesting in differences in progression, metastasis and outcomes among patients. Differences in cancer subtype, defined by the presence or absence of progesterone receptors (PR+) estrogen receptors (ER+) and human epidermal growth factor receptor 2 (HER2+) on the cancer cells, also contribute to differences in metastatic potential [5]. This heterogeneity extends to the cellular level in a single patient, where individual cancer and stromal cells in the tumor microenvironment can have significantly different impacts on cancer growth and metastasis [6]. Unravelling this web of connections between cancer and stromal cells will be critical for developing safe, effective and personalized treatments for cancer metastasis.

2.1.2 Tumor associated macrophages promote tumor growth and metastasis

Macrophages are the most abundant immune cells in mammary tumors and contribute to tumor development and metastasis by secreting inflammatory factors, remodeling the tumor microenvironment and degrading the extracellular matrix of the basement membrane [7-11]. TAMs facilitate tumor proliferation and angiogenesis through

continual extracellular matrix remodeling, while also suppressing adaptive immune responses targeted at tumor cells [12-14]. Macrophages and tumor cells secrete a variety of proteases including matrix metalloproteinases (MMPs), aspartyl proteases and cysteine cathepsins into the tumor microenvironment where they promote tumor angiogenesis through the cleavage of ECM into angiogenic growth factors, and allow tumor cells to enter the vasculature by compromising the basement membrane and degrading cellular adhesion molecules opening gaps for cancer cells to spread [8, 15].

2.2 Cysteine cathepsins

2.2.1 Cathepsins are highly regulated proteases

Cysteine cathepsins are potent degraders of extracellular matrix proteins involved in a host of pathologies including cancer, atherosclerosis and osteoporosis. There are 11 cysteine cathepsins (B, C, F, H, K, L, V, O, S, W, X/Z) including the most potent known human elastase, cathepsin V, and collagenase, Cathepsin K [16-19]. Cathepsins are components of a complicated regulatory proteolytic network that includes multiple biomechanical and biomechanical pathways, endogenous protease inhibitors and other proteases capable of activating pro-proteases into their mature active forms, or binding and degrading pro or mature proteases [20-24]. Many of these enzymes are implicated in extracellular matrix remodeling, and recent research suggests they are also involved in regulatory processes including apoptosis, autophagy, protein activation and antigen presentation [16, 25-28]. Cysteine cathepsin activity is also regulated by pH and redox conditions, with the majority of cathepsins degrading substrate most effectively in low pH and mildly reducing environments similar to lysosomes, with the notable exception of cathepsin S, which retains its activity at neutral pH [29].

2.2.2 Cathepsins promote cancer metastasis

Several cathepsins have been identified in tumor associated macrophages (B, C, K, L, S, V, X) and breast cancer cells (B, F, H, K, L, S, V, X) [25, 30]. Cathepsin K has been identified in both primary breast tumor sites as well as in bone metastatic sites [31]. Treatment with the pharmaceutical cathepsin K inhibitor odanacatib was found to significantly reduce metastatic bone disease in women with breast cancer compared to placebo [32]. Cathepsin V (also known as cathepsin L2) has been identified as an antigen using sera from patients with breast cancer [33]. Interestingly, while cathepsin L has been identified as a promoter of tumor growth, invasion and metastasis in breast cancer [34], loss of cathepsin L has been shown to promote intestinal neoplasia [35] and epidermal carcinoma [36]. Cathepsin B has a near ubiquitous tissue distribution, but it is overexpressed in many cancers including breast, cervix and prostate [37, 38]. Pharmacological inhibition of cathepsin B has been shown to decrease bone metastases in a mouse model of metastatic breast cancer [39]. Inhibition with a broad spectrum cysteine cathepsin inhibitor failed to suppress metastasis in a murine model of breast cancer, while a cathepsin B specific treatment significantly reduced secondary tumor volume in the spine and lung, suggesting some cathepsins may play anti-tumorigenic roles [39]. Cathepsins S and B produced by tumor associated macrophages have been shown to protect tumors from chemotherapy induced cell death in a murine model of breast cancer [40].

2.2.3 Pharmacological cathepsin inhibitors

Proteases have long been implicated in the development and metastasis of cancer. Several protease families of interest have been identified including matrix metalloproteinases (MMPs), and serine, aspartic and cysteine proteases. While gene

knockout, RNA interference and enzyme inhibitor treatments have shown efficacy in human *in vitro* and murine *in vivo* systems, successful implementation in human clinical trials has been curtailed by a variety of off-target effects including skin lesions and cardiovascular complications [41].

Upregulation of cathepsins in pathologies including cancer, cardiovascular disease and osteoporosis have made these enzymes attractive targets for pharmaceutical inhibition [42-44]. Multiple potent, specific small molecule inhibitors of different cathepsins have been developed, but despite showing clinical efficacy in trials for osteoporosis and bone metastasis, to-date no cathepsin inhibitors have been cleared for clinical use for any pathology due to unexpected side effects [32, 41, 45, 46].

Cathepsin inhibitors have a history of provoking unexpected results *in vivo* and *in vitro* [47, 48]. Healthy subjects treated with a specific cathepsin S inhibitor exhibited a dose-dependent increase in both total plasma cathepsin S and cathepsin S activity [49]. Our group has previously shown that treatment of MDA-MB-231 cells with the small-molecule cathepsin inhibitor E64 (*trans*-Epoxysuccinyl-L-leucylamido(4-guanidino)butane) can increase the amount of active cathepsin S, while decreasing the amount of active cathepsin L in breast cancer cells [50]. The underlying mechanism responsible for this unexpected reaction to inhibitor treatment has not been previously established. Developing experimental and computational tools to study the dynamic proteolytic regulatory network responsible for generating these unintuitive results is critical for the future development of safe and effective cathepsin inhibition strategies for the prevention of cancer metastasis and other pathologies.

2.3 Biological Challenges of Computationally Modeling Cathepsin Dynamics

2.3.1 Cathepsin "lifecycle"

In most cell types cysteine cathepsins are translated as preprocathepsins in the rough endoplasmic reticulum (ER) (Figure 2-1). Signal peptides target them to the endoplasmic reticulum lumen where the signal peptides are cleaved by signal peptidases, yielding procathepsins. Procathepsins are then post-translationally modified by glycosylation with high-mannose glycans. Procathepsins then pass through the Golgi and enter the trans-Golgi network where mannose residues are phosphorylated yielding mannose 6-phosphate (M6P) tags that drive their sorting into vesicles bound for late endosomes. Procathepsins are then activated to their mature forms in low pH environments, through the removal of their propeptide either autocatalytically, by other cathepsins, or by legumain/asparaginyl lysosomal endopeptidases as the late endosomes mature and fuse with lysosomes [16, 51-53]. In addition to this canonical pathway, under disease states or cathepsin overexpression conditions, cathepsins can overrun the Golgi pathway, bypassing M6P sorting and leading to secretion of inactive procathepsins [54, 55]. Macrophages are also known to secrete lysosomal enzymes under normal culture conditions and in response to stimuli such as hypoxia [56, 57].

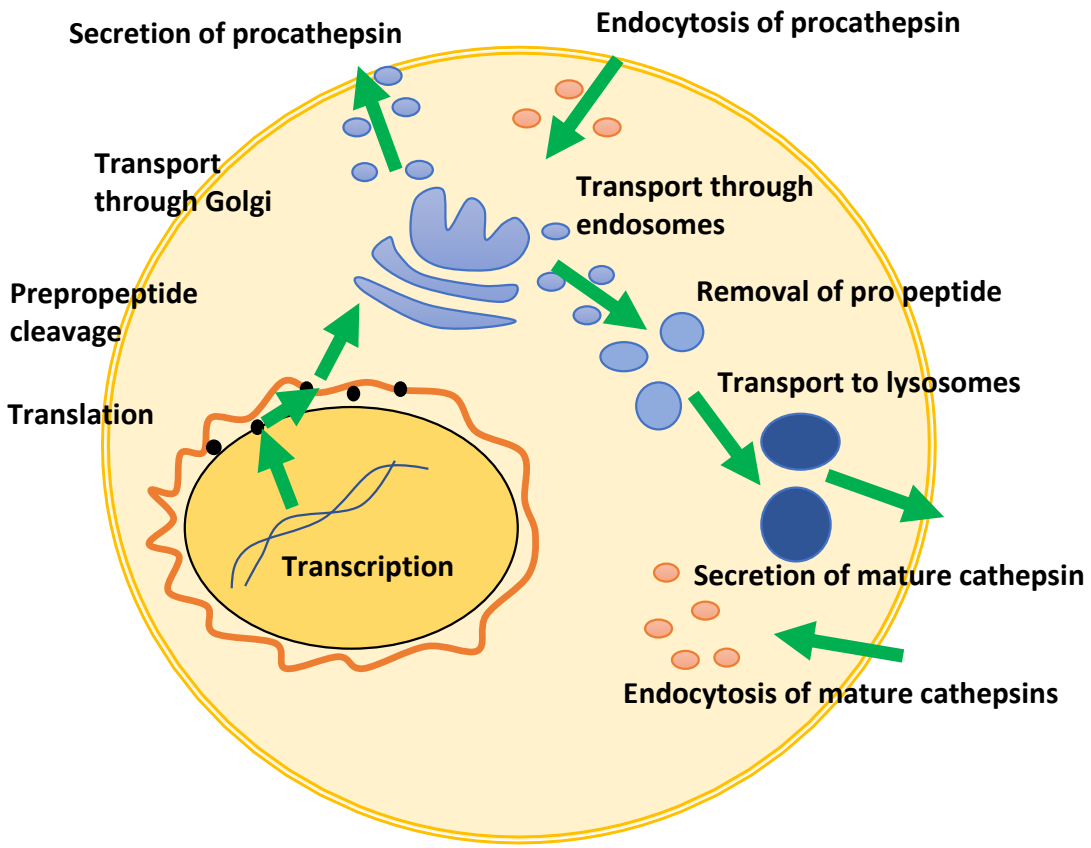

Figure 2-1: Overview of the cysteine cathepsin lifecycle in mammalian cells. Cathepsins are translated in the endoplasmic reticulum as inactive zymogens known as preprocathepsins where they are targeted to the ER lumen by leader sequences. The prepropeptide is cleaved and the inactive proenzymes are transported to the Golgi apparatus after proper protein folding. In the Golgi, procathepsins are targeted for secretion or transport through endolysosomal vesicles by glycosylation and mannose-6-phosphorylation. Procathepsins are activated to their mature forms by cleavage of the propeptide accomplished through a combination of autocatalytic or protease-mediated hydrolysis accelerated by the increasingly acidic and reducing environment of maturing endosomes. In some cells, mature cathepsins can be transported from endosomes and lysosomes to the plasma membrane for secretion, or entire endosomes and lysosomes can fuse with the plasma membrane releasing their contents to the extracellular space. In some cells, extracellular procathepsins and mature cathepsins can be endocytosed leading to intracellular cathepsin accumulation.

2.3.2 Cathepsin trafficking

Protein trafficking is an important consideration for cathepsin mathematical and computational modeling because the activity of cathepsins are tied to the pH and redox

conditions of their microenvironment. Procathepsins activate to their mature forms by cleavage of the propeptide under acidic conditions. This reaction can proceed autocatalytically or when catalyzed by another protease. Catalysis by ancillary proteases is also dependent on the environment, as the pH, redox, and presence of proteases, inhibitors and substrates all affect the activation process.

Secretion of inactive procathepsins or active mature cathepsins into the extracellular environment can drastically alter the proteolytic dynamics surrounding the cell, as substrates are degraded, endogenous inhibitors bind to the proteases and proteases become inactivated or degraded. Cathepsin inactivation, distinct from inhibition which requires an inhibitor, is not currently well understood, but is thought to encompass denaturation of the protein as well as oxidation of the active site cysteine and possibly some other currently unknown mechanisms including degradation by other proteases, a hypothesis postulated by our research team. Mature cathepsins can inactivate, particularly in the absence of substrates or inhibitors, resulting in enzyme that has not been degraded, but is unable to degrade other proteins.

2.3.3 Limitations of cathepsin animal models

Animal models have been instrumental to cathepsin research, establishing links between cathepsins and atherosclerosis, vascular repair, antigen presentation, cancer metastasis and more [58-61]. However, some fundamental biological differences between human and non-human proteases make cathepsin animal models challenging. While mouse cells express cathepsins B, L, S and K, they do not express cathepsin V. Confusingly mouse cathepsin L is the homolog to human cathepsin V, since it shares greater sequence identity with human cathepsin V than with human cathepsin L [62]. The case of odanacatib, a

cathepsin K inhibitor that underwent clinical trials for the treatment of osteoporosis and cancer metastasis, highlights another limitation of animal models. While odanacatib is an effective inhibitor of human cathepsin K, with a Ki of 0.18 ± 0.06 nM, it is much less effective at inhibiting mouse cathepsin K, with a Ki of 32.7 ± 6.8 nM, resulting in a 182 fold lower effectiveness of inhibition of mouse cathepsin K compared to human cathepsin K despite 86% sequence similarity between the two enzymes [63]. Odanacatib is likely not unique as a species-specific cathepsin inhibitor. The development and validation of better human cathepsin mathematical and computational models would not only reduce time and cost compared to using animal models, but increase the likelihood of identifying effective inhibitors for the treatment of human disease.

2.4 Approaches to computationally model cathepsin activity

In vitro and *in vivo* experiments on cathepsin dynamics are the gold standard, but these studies can be extremely costly and time consuming. Computational and mathematical modeling based on, and validated using, experimental data provides an alternative approach to hypothesis generation and testing necessary to unravel this complicated network of interactions incorporating the lifetimes and potency of this family of enzymes.

2.4.1 Substrate degradation and enzyme kinetics

The earliest uses of mathematical and computational modeling in cathepsin research relied on Michaelis-Menten (MM) principles of enzyme kinetics. Since then, more advanced modeling techniques have been employed on many different cathepsin family members. While substrate degradation by cathepsins can be represented with MM models, cathepsin activation and inactivation necessitate additional mathematical terms. Enzyme

kinetics and inhibition of different cathepsins have been described on different natural and artificial substrates, including elastin [64].

Mathematical modeling has been fundamental to discovering the complex role of ancillary molecules and inhibitors in cathepsin activity. Kinetic modeling of cathepsin K was used to describe the mechanisms by which glycosaminoglycans (GAGs) promote cathepsin K degradation of collagen [65]. Kinetic modeling has been used to investigate the poorly understood mechanisms related to cathepsin K inactivation and it has identified clusterin as a 'liberator' of cathepsin K, a molecule that protects an enzyme from inhibition without affecting the enzyme's activity [66]. Kinetic modeling of inhibition of cathepsin B by cystatin C identified a two-step inhibition process, involving the occluding loop of cystatin B, responsible for cathepsin B's observed resistance to cystatin C inhibition compared to other closely related cysteine cathepsins [67].

2.4.2 Ordinary differential equation modeling

Building on the foundation of work laid by kinetic modeling, mechanistic ordinary differential equation (ODE) modeling based on the law of mass action has been used to discover new mechanisms of cathepsin regulation. Several properties of these models make them attractive to cathepsin research, including their ability to incorporate dynamics, model continuous solutions and include multiple compartments. ODE models are able to simulate enzyme dynamics with different model species changing over time. This is useful when studying cathepsins because time sensitive reactions, including activation, removal of the propeptide, and inactivation, denaturation or oxidation of active site residues, can modify the ability of cathepsins to degrade substrates in a time dependent manner. ODE models can be used to generate continuous solutions, which allow for interpolation between

discrete data points. Utilizing multiple ODEs allows for the modeling of multi-compartment systems, allowing ODE models to incorporate movement of model species, such as secretion from intracellular to extracellular compartments.

ODE modeling is a well characterized and flexible tool that has been used extensively in systems biology to analyze metabolism, genetic regulation, cellular signaling and many other biological processes [68]. Applying this technique to cathepsins allows us to build on standardized methods and tools to formulate model diagrams, convert them into equations, estimate parameters and analyze the system under different conditions. Depending on the exact formulation of the model equations, the parameters attached to mechanistic ODE models can closely correspond to biologically observed reaction rates, allowing for relatively straightforward insights into the biological system being modeled (e.g. secretion rate in the model directly predicts how quickly cells are secreting an enzyme). This is in contrast to more abstract model formulations, where the model is able to fit the experimental data, but the parameters do not correspond directly components of the biological system.

The main limitations of ODE modeling are the assumptions inherent to the law of mass action and difficulties encountered in parameter estimation. The law of mass action assumes species in a compartment are completely mixed and able to freely move and interact with other species in the compartment. These assumptions are rarely completely satisfied in cellular systems, but well parameterized ODE models offer reasonable approximations of these systems while requiring significantly less computational resources than more intensive methods such as reaction-diffusion modeling. Determining the best parameters for an ODE model system based on experimental data is currently an unsolved

problem that is often addressed on a model-to-model basis. Computational optimization algorithms capable of dealing with large amounts of data have taken great strides in recent years, but identifying reasonable parameters is still a time consuming, iterative process. Generating and evaluating these potential solutions is a critical part of creating an ODE model capable of answering biological questions. Keeping these limitations in mind when constructing the initial model can have a major impact on the ability to identify reasonable parameters, since parameter estimation complexity increases exponentially as more equations and parameters are incorporated into the model.

ODE modeling has been used to characterize several systems involving cathepsins, substrates and inhibitors. Dynamic interactions between cathepsin B, cystatin C and β-amyloid aggregation were explored using a mechanistic ODE model, where cathepsin B was found to degrade β-amyloid plaques, cystatin C inhibited cathepsin B, but cystatin C also bound to β-amyloid precursors [69]. Mechanistic ODE modeling has elucidated cathepsin-on-cathepsin interactions between cathepsin S and cathepsin K, termed "cathepsin cannibalism" [70]. Further work has shown that these cannibalistic interactions are not limited to cathepsins S and K, but encompass cathepsins K, L, S and V with more interactions likely to exist (Ferrall-Fairbanks et al. in preparation). The existence of these cannibalistic interactions is hypothesized to control substrate degradation both through degradation of mature cathepsins, but also through cathepsins acting as "distractions" or substrates for other active cathepsins (Ferrall-Fairbanks et al. in preparation). Mechanistic ODE modeling also has been used to investigate an observed increase in active cathepsin S in cancer cells following treatment with a broad spectrum cathepsin inhibitor [50]. Model results indicated that the irreversible, covalently binding cathepsin inhibitor E64 could

become unbound from cathepsins under non-reducing SDS-page conditions, suggesting the inhibitor might not irreversibly bind the active site cysteine of cathepsins (Shockey et al. under review).

2.4.3 Protein structure modeling

Since cathepsins are attractive drug targets for cancer and other diseases such as osteoporosis and atherosclerosis, many groups are using computational methods to screen and identify potential cathepsin inhibitors. High throughput molecular docking successfully identified an allosteric small-molecule inhibitor of cathepsin K, several potentially irreversible binding inhibitors of cathepsin B and inhibitors of cathepsin L [71-73]. A cathepsin S structural model was generated using the X-ray crystal structure of cathepsin K, then DOCK software was used to identify a highly potent, nonpeptidic noncovalent cathepsin S inhibitor that is highly selective for human cathepsin S compared to mouse, dog, monkey and bovine cathepsin S or other closely related human enzymes including cathepsins L, F and K [74]. Three-dimensional quantitative structure-activity relationship modeling has been used in combination with ligand based pharmacophore modeling to identify novel cathepsin L inhibitors from a library of natural compounds [75]. A combination of quantum mechanical and molecular mechanical modeling has been used to investigate the molecular steps required for cathepsin K cleavage of a peptide substrate, identifying the acylation step as rate limiting for proteolysis [76].

Protein structure modeling has also been used to identify substrates and interactions with ancillary molecules other than inhibitors. A study on fibrinogen used a combination of molecular docking and protein sequence modeling to predict cleavage sites by cathepsins K, L and S on fibrin and fibrinogen molecules where cathepsins K, L and S

were experimentally confirmed to degrade fibrin [77]. Molecular docking has been used to identify allosteric binding sites for heparin on cathepsin B and elucidating the mechanism responsible for heparin protecting cathepsin B from inactivation at alkaline pH [78].

Electrostatic modeling of cathepsins at different pHs has also been employed to study cathepsin cannibalism. At acidic pH the electrostatics of cathepsins S and K are less favorable for binding between the two enzymes, but as the pH becomes neutral the electrostatic model predicts more favorable interactions between cathepsins S and K (Kieslich et al. in preparation). This increase in electrostatic favorability indicates cathepsin cannibalism could be more likely at neutral pH than acidic pH, suggesting cathepsin cannibalism could be a natural mechanism for degrading cathepsins that have escaped their acidic pH niches.

2.4.4 Statistical modeling

Statistical modeling has seen relatively little use compared to kinetic and structural analyses, but the predictive power and flexibility of these models make them useful in non-mechanistic applications. Statistical modeling using cathepsin expression and kinase activation has been used to predict severity of cancer progression and invasiveness, identifying patient-specific monocyte phenotypes predictive of tumor invasiveness, independent of the cancer cells themselves [79]. Statistical modeling is also useful for analysis of cathepsin inhibitor drug trials and designing dosing regimens. Analysis of one cathepsin K clinical trial could predict the bone mineral density following different dosages of a cathepsin K inhibitor [80].

2.5 Conclusions

Cysteine cathepsins are challenging enzymes to study with traditional molecular biology techniques. Cathepsin activity is largely regulated post-translationally by a dynamic network of proteases, substrates, inhibitors and ancillary molecules. Mathematical and computational modeling offer new methods for generating and testing hypotheses in protease biology. These techniques have already proved indispensable in understanding mechanisms including cathepsin cannibalism and allosteric cathepsin regulation as well as the identification of novel pharmaceutical cathepsin inhibitors.

3 CHARACTERIZING THE PROTEOLYTIC PROFILES OF HUMAN MACROPHAGES

3.1 Introduction

Monocytes and macrophages are key components of the adaptive immune system and provide a first line of defense against infection and other pathologies. Circulating monocytes extravasate from the blood to sites of injury and inflammation where they differentiate into macrophages in response to environmental stimuli including a variety of inflammatory cytokines [81]. While macrophages are extremely effective at phagocytosing foreign pathogens and stimulating adaptive immune responses through antigen presentation, macrophages can actually exacerbate serious pathologies including atherosclerosis and multiple types of cancer. Proteases, particularly cysteine cathepsins, secreted by macrophages can weaken extracellular matrix barriers, facilitating the expansion of smooth muscle cells into the vascular intima or the extravasation of cancer cells into the bloodstream leading to cancer metastasis [9, 82]. Characterization of monocyte and macrophage proteolytic profiles could provide insights into the mechanisms of these pathologies and inform the development of future treatments, particularly cathepsin inhibitors.

In general, computational modeling of biological systems benefits from absolute quantification of biological variables, particularly protein levels. However, many biological questions can be answered with experiments that do not quantify variables in an absolute sense. This chapter details attempts to quantify the cathepsin profiles of human macrophages by measuring active cathepsins with cathepsin zymography and levels of pro and mature cathepsin by Western blot. These data are necessary to parameterize models of

the proteolytic interactions between tumor associated macrophages and cancer epithelial cells, since both these cell types are known to produce and secrete multiple proteases into the shared milieu of the tumor microenvironment. Isolating the proteolytic contributions of macrophages will allow us to generate personalized tumor models, due to the documented variance in cathepsin profiles of macrophages from different people [79, 83]. We also document an unexpected response to cathepsin inhibitor treatment, where macrophages treated with a cathepsin inhibitor have *increased* active and pro cathepsins with no measurable difference in mature protease.

3.2 Materials and Methods

3.2.1 Thp-1 monocyte and macrophage culture and sample collection

Undifferentiated Thp-1 monocytes were cultured in suspension in RPMI 1640 medium (Lonza) supplemented with 10% FBS, 1% L-glutamine and 1% Penicillin/Streptomycin. Thp-1 monocytes were serially diluted and different numbers of cells were lysed in 50 µL of zymography lysis buffer. For each cell count, 10 µL of lysate were analyzed for active proteases by cathepsin zymography. Thp-1 monocytes were differentiated into Thp-1 macrophages through treatment with 100 nM phorbol 12-myristate 13-acetate (PMA) for 24 hours. Following differentiation, macrophages were cultured in media containing 10µg/mL cycloheximide for 24 hours to stop protein translation. Cells were switched to serum free media after 24 CHX treatment to begin time course measuring secreted cathepsins. At each time point, cell media was collected and acetone precipitated by mixing 5:1 acetone to media, holding on ice for 30 minutes, then spinning at 3000g for 10 min, before aspirating acetone and resuspending protein precipitate in 100 µL zymography lysis buffer for a final concentration of 10x.

Macrophages were lysed in plate with zymography lysis buffer. Macrophages treated with E64 were differentiated with PMA as described above then incubated with different concentrations of E64 for 24 hours.

3.2.2 Multiplex cathepsin zymography

Equal volumes of Thp-1 monocyte lysate, equal volumes of Thp-1 macrophage media and equal protein amounts of Thp-1 macrophage lysate, were mixed with nonreducing SDS-PAGE loading buffer and cathepsin zymography was performed as previously described [84, 85]. Briefly, samples were loaded into a gelatin-embedded polyacrylamide, lysate proteins were separated by molecular weight during electrophoresis. Following electrophoresis, cathepsins were renatured in polyacrylamide gels before being incubated overnight in cathepsin-specific assay buffer at 37°C. The gels were then stained with Coomassie Blue, then destained. White bands indicate the presence of active cathepsins capable of degrading gelatin embedded in the polyacrylamide gel. Total protein concentration of cell lysate was measured using the Pierce Micro BCA Protein Assay (Thermo Scientific).

3.2.3 Western blot

Total protein concentration in cell lysate was determined by BCA, then equal amounts of protein were mixed with reducing loading dye, briefly boiled, then loaded into 12.5% polyacrylamide gels and separated by electrophoresis. Proteins were transferred onto nitrocellulose membranes, blocked and probed with primary antibodies for human cathepsin S (R&D Systems) and human cathepsin L (R&D Systems). Membranes were probed with secondary antibodies (LI-COR) then imaged on a LI-COR Odyssey CLx Imaging System.

3.2.4 Statistical analysis

Western blot and zymography band intensities were quantified by densitometry in ImageJ and data were normalized by dividing each band intensity on a gel by the intensity measured from cells treated with PMA and 0 μM E64 (setting this group to 1 across all gels). Two-tailed Student's t-test with two-sample equal variance was performed on all statistical analysis.

3.3 Results

3.3.1 Active cathepsin V detectable in 1.5 μg of protein in Thp-1 monocyte lysate

Cathepsin zymography is a useful method for measuring active cathepsins in a variety of cell and tissue samples. Assessing zymography results from homogenized tissues like breast tumors requires validation for the expected signal and limits of detection of different types of cells that could be contributing proteases to the tissue. Active cathepsins were detectable across all cell counts down to 100,000 cells, with 50,000 cells not producing a detectable zymography signal (Figure 3-1A). Total protein content of the lysate samples was quantified by bicinchoninic acid (BCA) assay and plotted against cell count (Figure 3-1B). Assuming complete lysing of Thp-1 monocytes, 20% of cell lysate was analyzed by cathepsin zymography indicating active cathepsins can be detected in the lysate of as few as 20,000 cells. The measured total protein concentration of the 100,000 cell lysate was 0.1534 μg/μL, of which 10 μL were loaded for zymography, indicating active cathepsins can be detected in as little as 1.5 μg of protein from Thp-1 monocyte lysate.

Comparison to a recombinant cathepsin V standard curve was used to determine the amount of intracellular active cathepsin V in Thp-1 monocytes. Serially diluted recombinant cathepsin V and Thp-1 monocytes were loaded and analyzed by cathepsin zymography (Figure 3-2A). Results were quantified by densitometry and the standard curve of known amounts of active cathepsin V were used to calculate the active cathepsin V content of the different numbers of Thp-1 monocytes (Figure 3-2B). The calculated active cathepsin V content of a single Thp-1 macrophage is 53.6 pg. The reported volume of human monocytes is between 46.8 to 78.38 μm^3 [86]. The estimated volume of lysosomes, where cathepsin V is assumed to reside, is 2.5% of cell volume in human macrophages, corresponding or 1.17 to 1.96 μm^3 [87]. Assuming monocytes have a similar relative lysosomal volume compared to macrophages yields an estimated active cathepsin V concentration of between 1.1 to 1.8M in lysosomes of Thp-1 monocytes.

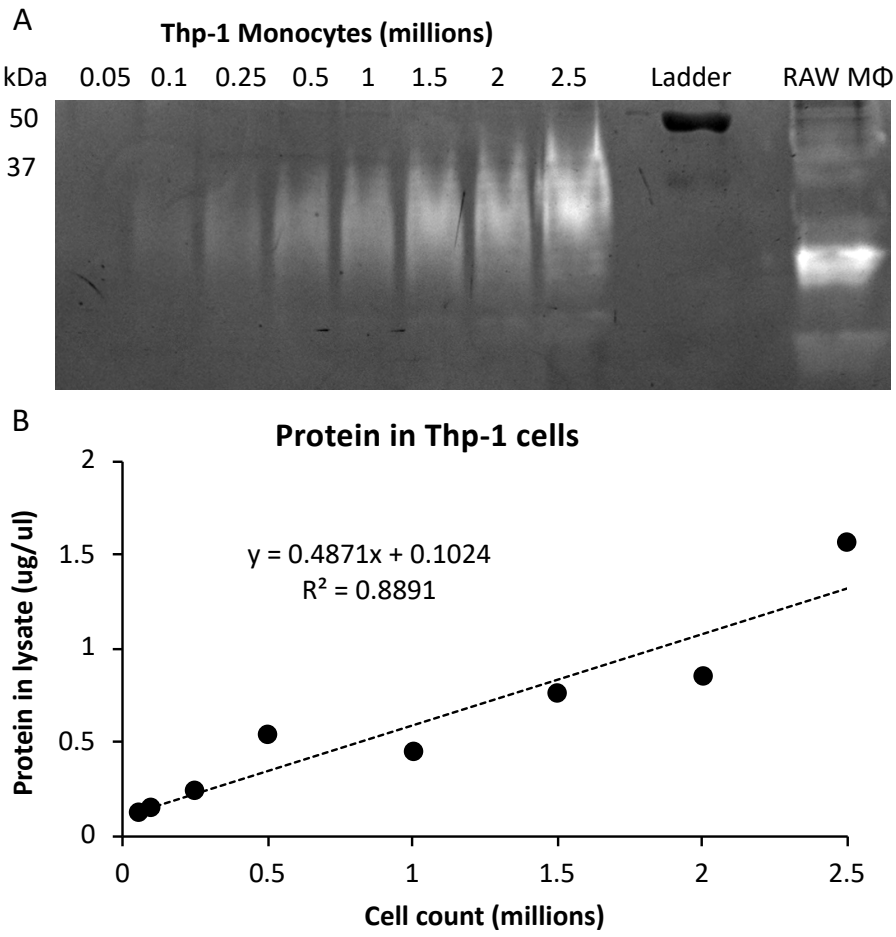

Figure 3-1: Active cathepsins from 100,000 Thp-1 monocytes detectable by cathepsin zymography. Thp-1 monocytes were serially diluted and lysed in 50 µL cathepsin zymography buffer. Active cathepsins were detectable at all cell counts greater than 50,000 cells (A). Total protein concentration of Thp-1 lysate was quantified by BCA. Protein content increased linearly with the number of cells in lysate (B).

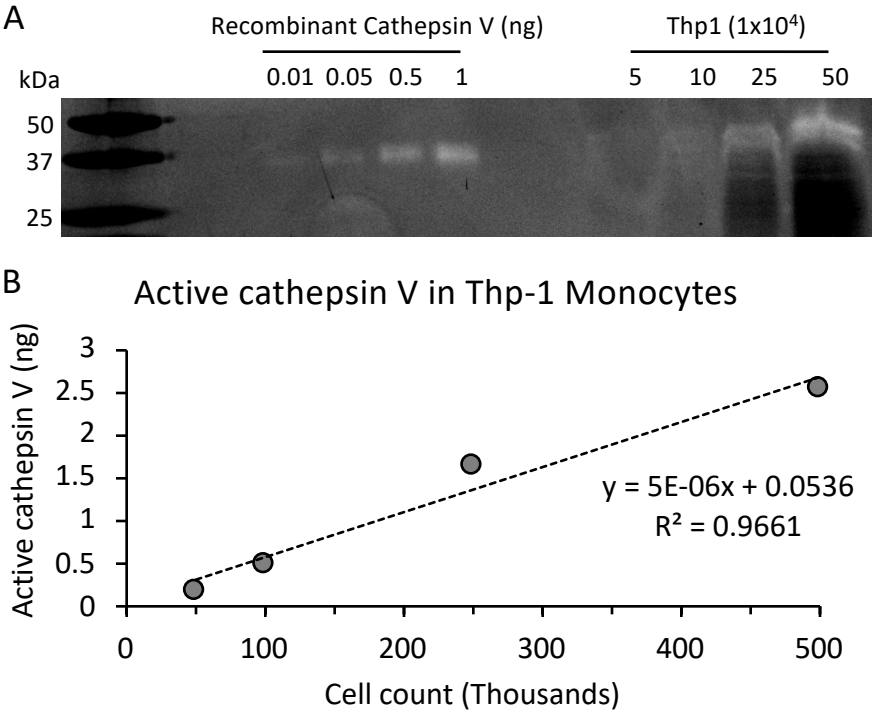

Figure 3-2: Active cathepsin V content of Thp-1 monocytes. Thp-1 monocytes were serially diluted and lysed in 50 µL cathepsin zymography buffer, and equal volumes of Thp-1 lysate and a recombinant cathepsin V standard curve were analyzed by cathepsin zymography (A). Active cathepsins were quantified by densitometry and active cathepsin V content of Thp-1 monocytes was calculated by comparison with the recombinant cathepsin V standard curve.

3.3.2 Quantifying dynamics of cathepsins L and S in Thp-1 macrophages

Macrophages, particularly macrophages in tumors and other sites of inflammation, are known to produce and secrete a variety of proteases, but the precise rates of production and secretion of these proteases have not been well characterized [88, 89]. Intracellular protein levels result from a constant balance of transcription, translation, secretion and degradation. Measuring steady state levels of intracellular proteins is relatively straightforward but determining the rates of influx and efflux of specific proteins can be more challenging due to the variety of cellular processes constantly influencing protein concentration. Interrupting one of these processes can provide the perturbation of the

system necessary to better understand the balance of processes leading to cellular homeostasis.

To better characterize the rates of intra- and extracellular accumulation of cathepsins L and S in macrophages, we interrupted protein production with the translation inhibitor cycloheximide. Thp-1 monocytes were differentiated into macrophages by stimulation with phorbol 12-myristate 13-acetate (PMA), then incubated with cycloheximide (CHX) for 24 hours to halt protein synthesis. A time course of lysate and media was collected over the next 24 hours and expression of cathepsins L and S was measured by Western blot (Figure 3-3 & Figure 3-4). This experiment revealed cathepsins L and S have markedly different intracellular dynamics in Thp-1 macrophages. Intracellular pro- and mature cathepsin L was almost completely undetectable after 24 hours of CHX treatment, indicating that cathepsin L is quickly degraded or secreted. After removal of CHX, steady state levels of cathepsin L rebounded, pointing to high intracellular turnover of cathepsin L (Figure 3-3A). In contrast, levels of cathepsin S remained relatively constant after CHX treatment to the end of the 24-hour time course, suggesting cathepsin S has greater intracellular stability than cathepsin L, and is potentially produced, secreted and degraded more slowly than cathepsin L (Figure 3-3B).

A similar dynamic was observed extracellularly, where procathepsin L was secreted more quickly than procathepsin S but plateaued by 4 hours indicating that it is likely degraded extracellularly (Figure 3-4A). Mature cathepsin L appeared to reach steady state by 8 hours, accumulating more slowly than procathepsin L, likely due to both a slow secretion of mature cathepsin L and extracellular activation of procathepsin L to mature. Extracellular pro and mature cathepsin S accumulated at similar relative rates, increasing

linearly throughout the entire 24-hour time course suggesting a relatively slow rate of secretion coupled with a low rate of extracellular cathepsin S activation and degradation.

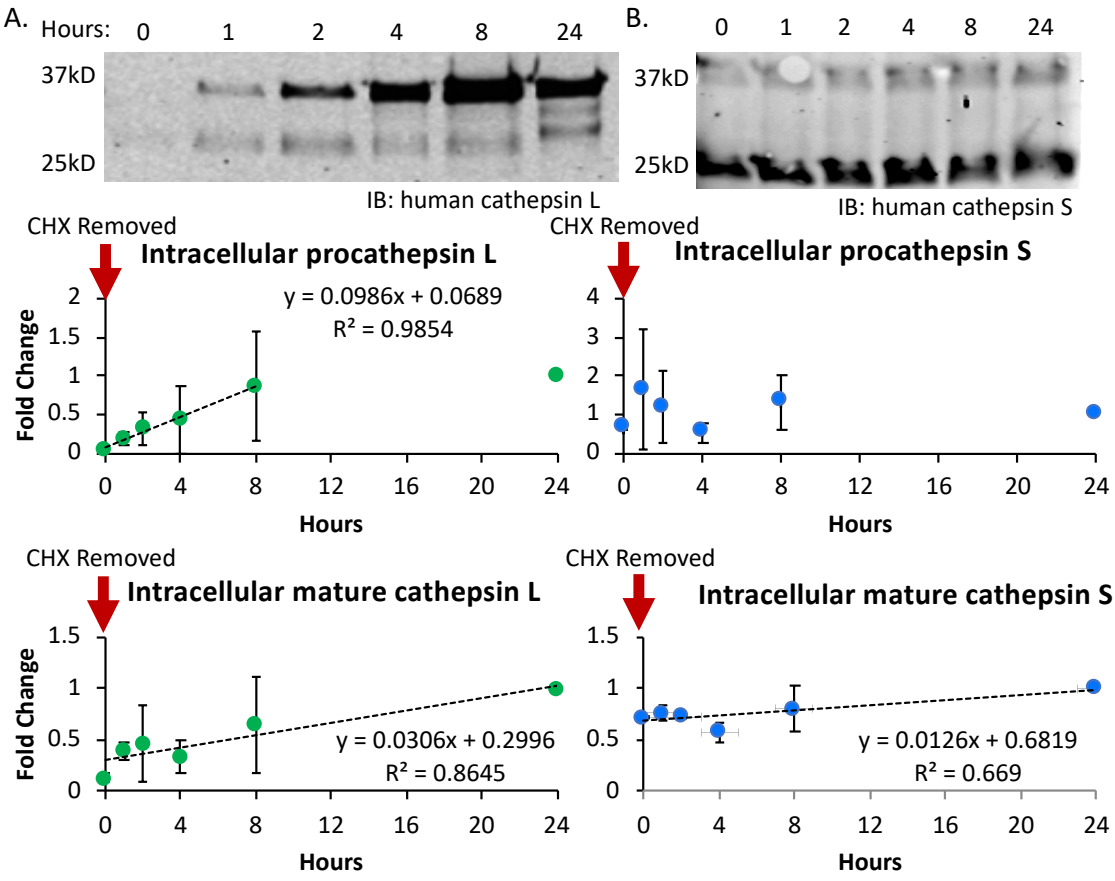

Figure 3-3: Translation inhibitor depletes cathepsin L but not cathepsin S in Thp-1 macrophages. Thp-1 monocytes were differentiated into macrophages by PMA treatment, then incubated with the translation inhibitor cycloheximide (CHX) for 24 hours. Cells were changed to fresh media and a time course of lysate was collected over 24 hours. Pro- and mature cathepsin L (A) and cathepsin S (B) content were measured by Western blot and quantified by densitometry. Intracellular pro- and mature cathepsin L was almost undetectable following 24 hours of CHX treatment. Procathepsin L increased linearly for the first 8 hours, plateauing before 24 hours. Mature cathepsin L increased linearly for the entire 24 hour time course. Procathepsin S was present after CHX treatment, only accumulating slightly over 24 hours. Mature cathepsin S was also detectable after CHX treatment, accumulating linearly over 24 hours, though much more slowly than cathepsin L (n=2).

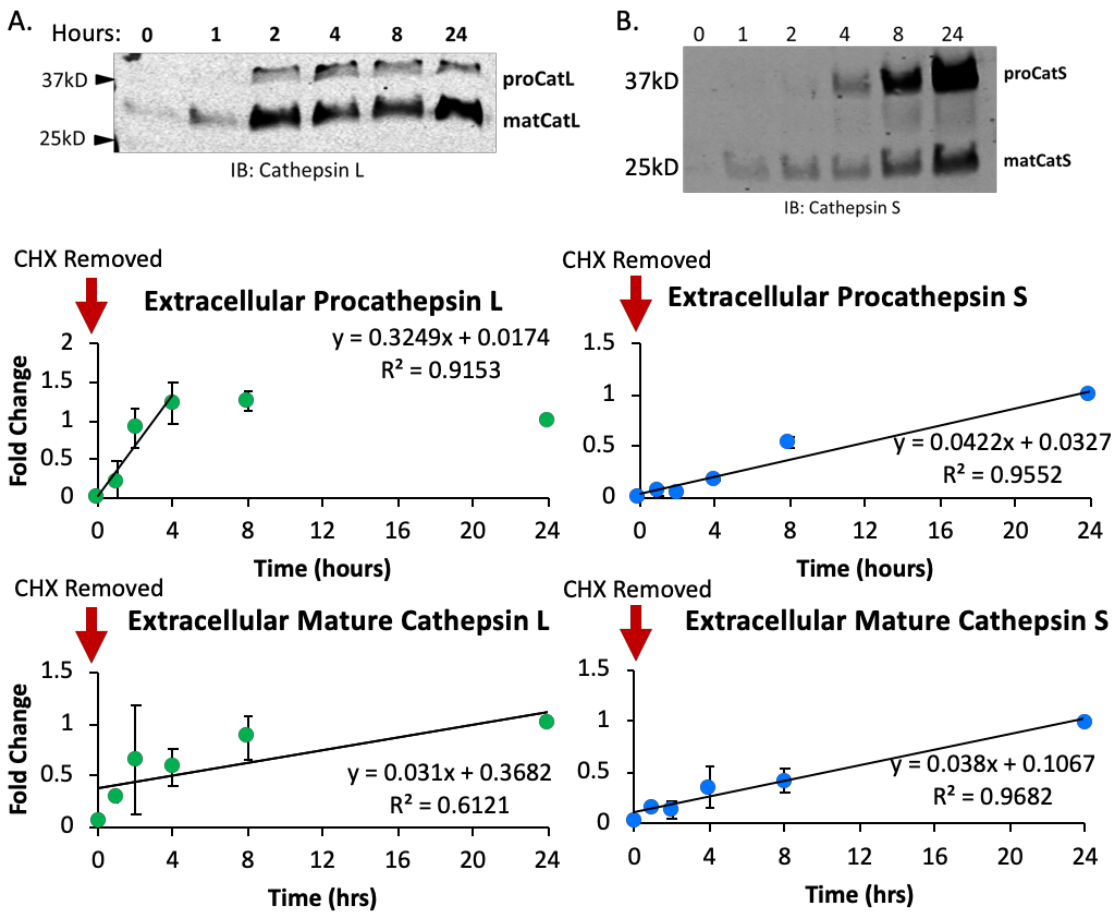

Figure 3-4: Cathepsin L accumulates more quickly than cathepsin S in Thp-1 media. Thp-1 monocytes were differentiated into macrophages by PMA treatment, then incubated with the translation inhibitor cycloheximide (CHX) for 24 hours. Cells were changed to fresh media and a time course of media was collected over 24 hours. Pro- and mature cathepsin L (A) and cathepsin S (B) content were measured by Western blot and quantified by densitometry. Pro and mature cathepsin L accumulated rapidly over the first 4 hours before plateauing by 8 hours, with procathepsin L declining slightly by 24 hours. Pro and mature cathepsin S accumulated linearly, but more slowly than cathepsin L (n=2).

3.3.3 Cathepsin inhibition stimulates increased active cathepsins in macrophages

Pharmaceutical small molecule cathepsin inhibitors are currently under investigation to treat a variety of pathologies including osteoporosis and cancer metastasis [44]. We decided to explore the impacts of cathepsin inhibition on the proteolytic profiles of macrophages by treating Thp-1 macrophages with the broad-spectrum cathepsin

inhibitor E64 and measuring active cathepsins by cathepsin zymography (Figure 3-5). We hypothesized that E64 would decrease active cathepsin levels in Thp-1 macrophages, because it is a broad spectrum irreversibly binding cathepsin inhibitor. Unexpectedly, treatment with the cathepsin inhibitor caused an increase in the amounts of active cathepsins V and S, while decreasing the amount of active cathepsin L in a dose dependent manner. Amounts of active V, S and L were increased in Thp-1 macrophages differentiated by PMA treatment, compared to Thp-1 monocytes, consistent with their differentiation towards a macrophage phenotype.

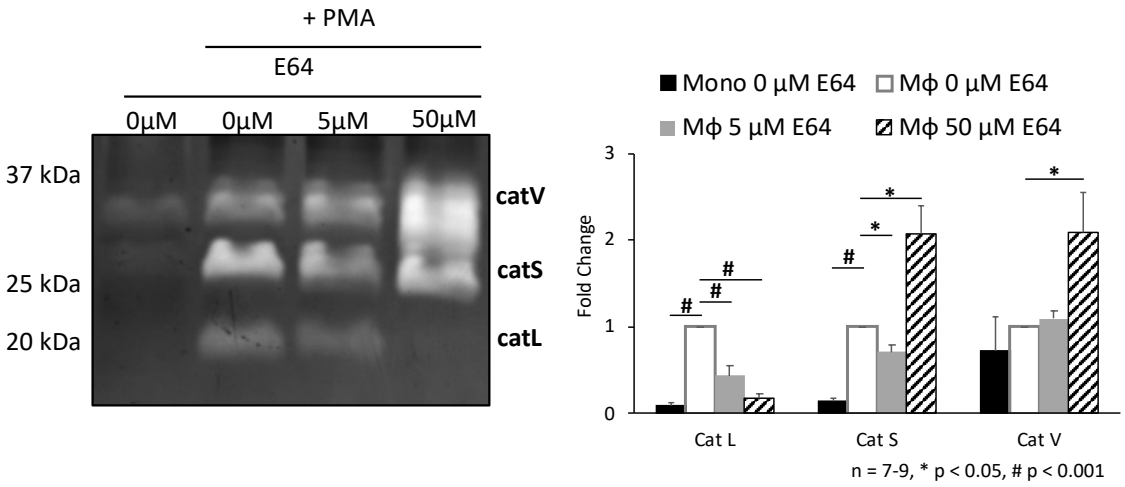

Figure 3-5: Macrophages have reduced amounts of active cathepsin L after incubation with E64, but higher amounts of active cathepsins S and V. Thp-1 monocytes were stimulated to differentiate into macrophages with PMA, and confirmed by the upregulated amounts of active cathepsins L, S, and V. These differentiated macrophages were then incubated with 0, 5, or 50 μM E64 for 24 hours, lysed, and equal amounts of protein were loaded for multiplex cathepsin zymography. E-64 incubation caused an 80% loss of active cathepsin L in the macrophages at 50 μM dose, but two-fold amounts of active cathepsin S and cathepsin V (n=7-9, * $p < 0.05$, # $p < 0.001$ by two-tailed t-test), as quantified and shown in the densitometry.

To determine if the measured increases in active cathepsins were due to increased protein expression, we measured cathepsin L and S levels in E64 treated Thp-1 macrophages by Western blot (Figure 3-6A). Pro- and mature cathepsin S and procathepsin

L were significantly increased after treatment with 50 µM E64 (Figure 3-6B). An increase in mature cathepsin S was consistent with the observed increase in active cathepsin S detectable by zymography. Pro- and mature cathepsins L and S were significantly increased after PMA stimulated differentiation.

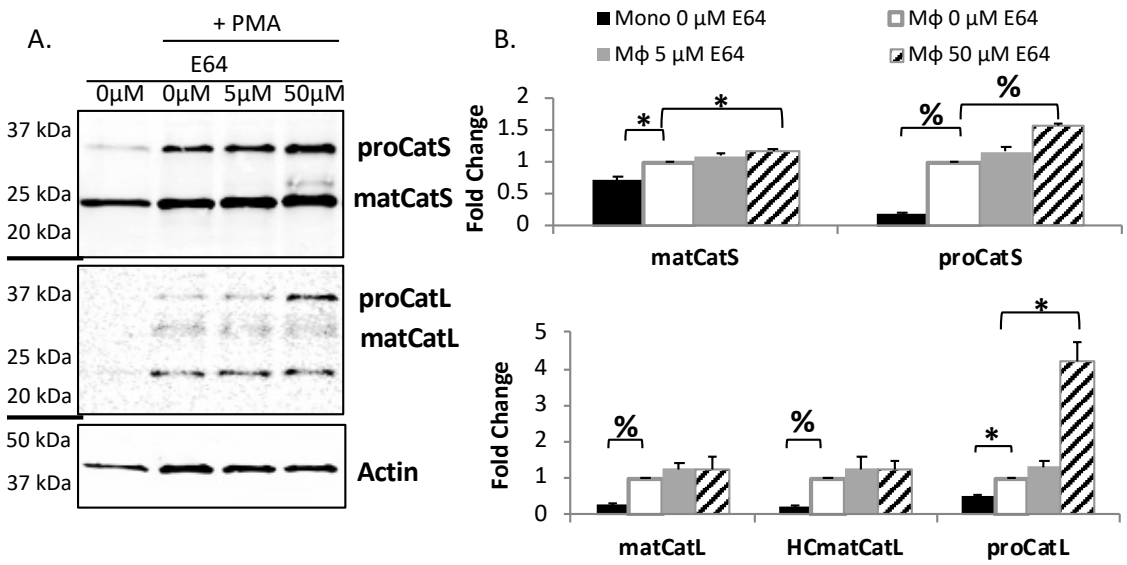

Figure 3-6: Cathepsin inhibitor treatment increases cathepsin L and S in Thp-1 macrophages.
Thp-1 monocytes were differentiated into macrophages by PMA treatment and incubated with the cathepsin inhibitor E64 for 24 hours. Cathepsin L and S levels were measured by Western blot. PMA treatment increased pro and mature cathepsins L and S in Thp-1 macrophages. Pro and mature cathepsin S and pro cathepsin L were increased in cells treated with 50 µM E64 (n=3 * $p < 0.05$, % $p < 0.01$ by two-tailed t-test).

3.4 Discussion

Different protein assays were used to quantify the cathepsin levels in human Thp-1 monocytes and macrophages. Active cathepsins, particularly cathepsin V, were detectable in as little as 1.5 µg of protein from Thp-1 lysate. The active cathepsin V content of a Thp-1 monocyte was calculated to be 53.6 pg, with a potential lysosomal concentration of between 1.1 to 1.8M. Intra- and extracellular dynamics of cathepsins S and L were shown to be quite distinct from each other in Thp-1 macrophages, with cathepsin L

production, secretion and turnover being significantly higher than cathepsin S. Finally, treatment with the cathepsin inhibitor E64 increased active cathepsins V and S in macrophages while decreasing active cathepsin L. E64 treatment caused a corresponding increase in pro and mature cathepsin S, consistent with their increased active levels measured by zymography. Interestingly, mature cathepsin L did not change in response to E64, despite a drop in zymography detectable active protein. Procathepsin L levels increased after E64 treatment, pointing to a reduced potential for intracellular activation of cathepsin L following E64 treatment consistent with cathepsin L or other cysteine proteases activating L in lysosomes.

CHX treatment for 24 hours almost completely depleted pro- and mature cathepsin L levels in Thp-1 macrophages. Interestingly, both intra- and extracellular levels of procathepsin L accumulated quickly and reached steady state levels 8 hours after CHX was removed. Intracellular cathepsin L reaching a steady state is not particularly surprising, since lysosome biogenesis, including the production of lysosomal proteases, is under transcriptional control and denatured cathepsin L can be degraded in lysosomes [90, 91]. Extracellular cathepsin L levels reaching steady state is more surprising, since they suggest some extracellular degradation of cathepsin L or a modulation in the production or secretion of cathepsin L by macrophages in response to concentration of cathepsin L or some other control mechanism. A likely scenario is an initial ramp-up in cathepsin L production and secretion post-CHX treatment, then some unknown regulatory process turns down synthesis of new enzyme, possibly cathepsin B controlling lysosomal biogenesis [92]. Notably, extracellular procathepsin L levels actually reach steady state by 4 hours, while intracellular procathepsin L continues to increase for an additional 4 hours

after the end of CHX treatment. Extracellular cathepsin L levels stabilizing before intracellular levels further supports the hypothesis of extracellular cathepsin L degradation since it is unlikely cathepsin L secretion was stopped entirely at an earlier time point, while intracellular cathepsins continued to accumulate and reached steady state hours later.

Active cathepsins V and S increasing in response to E64 treatment, while active cathepsin L decreases, is a surprising result since E64 is a broad spectrum cysteine cathepsin inhibitor that is thought to bind irreversibly to the cysteine in the active site of these enzymes [93]. An increase in procathepsin L following inhibitor treatment could point to cathepsin inhibition causing a decrease in reported autocatalytic activation of procathepsin L [94]. The increase in both pro- and mature cathepsin S suggests an increase in protein production, or decrease in protein secretion, rather than just a change in activation from pro to mature enzyme. Additional experiments will be required to determine any possible transcriptional regulation, or changes in enzyme secretion, impacted by E64 treatment. Additional experimental evidence will be required to investigate the effects E64 has on total cathepsin V protein levels as well. A similar response to E64 treatment by cathepsins L and S in breast cancer cells has been documented previously [50] and this analysis is extended to other breast cancer cell lines and primary patient tissue in Chapter 7.

4 CATHEPSIN DYNAMICS IN MULTICELLULAR SYSTEMS

4.1 Introduction

For many years, it was accepted that cathepsin activity was confined to specific intracellular compartments, particularly endolysosomal vesicles, where the acidic pH and reducing environment are amenable to cathepsins playing roles in antigen presentation and bulk protein turnover [91, 95]. More recent work has documented cathepsins in non-lysosomal intra and extracellular environments under healthy and pathologic conditions. Cathepsins have been identified in the cytoplasm, nucleus and mitochondria, as well as extracellularly where they are involved in extracellular matrix remodeling and cleavage of cell adhesion molecules [50, 96-98].

Following its discovery, cathepsin K was thought to be almost exclusively found in osteoclasts, where it plays a pivotal role in degrading type I collagen during bone resorption [99]. Since then, cathepsin K has been implicated in numerous cancers including breast and prostate [42, 100]. Cathepsin K is present both at primary tumor sites, as well as secondary metastatic sites particularly in the bone, and cathepsin K inhibition can decrease bone metastases among breast cancer patients [32]. Macrophages and other antigen presenting cells are known to produce and secrete cathepsin S, particularly under inflammatory conditions [101, 102]. Macrophages in bone metastatic sites of prostate cancer have also been shown to produce cathepsin K [60].

While cathepsins are potent degraders of extracellular matrix proteins, including collagen, elastin and laminin, cathepsin S has been shown to degrade cathepsin K *in vitro* even in the presence of matrix substrates [70]. Since cathepsin cannibalism was first discovered, a combination of recombinant enzyme experiments and computational

modeling techniques have identified likely cannibalistic interactions including: cathepsin S on cathepsin K, cathepsin K on cathepsin L, cathepsin L on cathepsin K and cathepsin K autodegradation [103]. Autodegradation has been observed previously in some proteases with wide substrate specificities, including human trypsin and bacterial thermolysin, but seems to be rarer among less promiscuous proteases [104, 105].

While cathepsin cannibalism has been measured *in vitro*, it has not been observed in cells. We sought to create a cellular system, where the effects of cathepsin cannibalism between cathepsins S and K could be observed. Data from these cellular experiments was used to parameterize a mathematical model capable of predicting the relative contributions of cathepsins S and K to cathepsin K degradation.

4.2 Materials and Methods

4.2.1 Fluorescently quenched elastin to measure cathepsin proteolysis

Five picomolar recombinant cathepsins S and K were incubated alone and in combination in serum-free DMEM containing 0.05mg/ml DQ elastin at 37°C. Fluorescence was measured at 10-minute intervals for 2 hours. Background fluorescence (measured with media containing DQ elastin control) was subtracted out of all reported measurements.

4.2.2 Generation of cathepsin overexpressing cell lines

HEK293T cells were stably transfected with plasmids coding for 6x-Histidne tagged wildtype and mutant cathepsins as previously described [103]. Cells were transfected with linearized cathepsin plasmids using lipofectamine 2000 (Invitrogen) and

then cultured in selection antibiotics for 4 weeks to ensure integration of plasmid DNA into the cell genome.

4.2.3 Coculture of cathepsin overexpressing HEK293T cells

HEK293T cells were cultured in DMEM (Lonza, supplemented with 10% FBS, 1% L-glutamine, 1% Penicillin/Streptomycin) and seeded in 6 well plates at 100,000 total cells per well overnight (e.g. 50,000 cathepsin K producing cells and 50,000 cathepsin S cells). Cells were washed with PBS then switched to DMEM without FBS. After 24 hours, media was collected and centrifugally concentrated ~10 fold in Vivaspin 500 concentrators (15,000g, 40 minutes, 4°C) before storing concentrated media at -80°C. Cells were lysed in zymography lysis buffer, sonicated and stored at -80°C.

4.2.4 Western blot

Total protein concentration in cell lysate and concentrated media was determined by BCA, then equal amounts of protein were mixed with reducing loading dye, briefly boiled, then loaded into 12.5% polyacrylamide gels and separated by electrophoresis. Proteins were transferred onto nitrocellulose membranes, blocked and probed with primary antibodies for human cathepsin K (Proteintech) and human cathepsin S (R&D Systems). Membranes were probed with secondary antibodies (LI-COR) then imaged on a LI-COR Odyssey CLx Imaging System.

4.2.5 Time course of cathepsin accumulation in media of cathepsin overexpressing cells for dynamic model fitting

HEK293T cells (cathepsin S producing cells alone, cathepsin K producing cells alone or cathepsin S and cathepsin K coculture) were seeded in 6 well plates at 100,000

total cells per well overnight. Cells were washed with PBS then switched to DMEM without FBS. Media was collected after 0, 2, 4, 6, 8 and 24 hours and concentrated by adding 5 mL of cold acetone per 1 mL of cell culture media, incubating on ice for 30 minutes, then collecting the protein precipitate by centrifugation (3000g, 10 minutes, 4°C) and resuspending in a volume of lysis buffer resulting in a 10 fold concentration. Amounts of pro- and mature cathepsins S and K were measured by Western blots loaded with standards of known amounts of recombinant cathepsins. Absolute amounts of cathepsins in the cell culture media were estimated by comparing sample bands to the recombinant standard curve and calculating molar concentrations in the original, unconcentrated conditioned media.

4.2.6 Systems biology modeling of cathepsin K and S producing cells

The following system of ordinary differential equations was developed to model the protein dynamics in the media of cathepsin overexpressing HEK293Ts based on the model diagram in Figure 4-3.

Model Equations:

$$\frac{dProS}{dt} = k_{S_{sec}} * Scells - k_{S_{act}}[ProS] \quad (1)$$

$$\frac{dMatS}{dt} = k_{S_{act}}[ProS] - k_{S_{\deg}}[MatS] - k_{SonK}[MatS][ProK] - k_{SonK}[MatS][MatK] + k_{SoffK}[MatS\_ProK] + k_{SoffK}[MatS\_MatK] + k_{ScatK}[MatS\_ProK] + k_{ScatK}[MatS\_MatK] \quad (2)$$

$$\frac{dProK}{dt} = k_{k_{sec}} * Kcells - k_{k_{act}}[ProK] - k_{SonK}[MatS][ProK] + k_{SoffK}[MatS\_ProK] - k_{KonK}[MatK][ProK] + k_{KoffK}[MatK\_ProK] \quad (3)$$

$$\frac{dMatK}{dt} = k_{k_{act}}[ProK] - k_{k_{\deg}}[MatK] - k_{SonK}[MatS][MatK] +$$
$$k_{SoffK}[MatS\_MatK] - 2 * k_{KonK}[MatK]^2 + 2 * k_{KoffK}[MatK\_MatK] +$$
$$k_{KcatK}[MatK\_MatK] - k_{KonK}[MatK\_MatK] + k_{KoffK}[MatK\_MatK] +$$
$$k_{KcatK}[MatK\_MatK] \quad (4)$$

$$\frac{dMatS\_ProK}{dt} = k_{SonK}[MatS][ProK] - k_{SoffK}[MatS\_ProK] - k_{ScatK}[MatS\_ProK] \quad (5)$$

$$\frac{dMatS\_MatK}{dt} = k_{SonK}[MatS][MatK] - k_{SoffK}[MatS\_MatK] - k_{ScatK}[MatS\_MatK] \quad (6)$$

$$\frac{dMatK\_ProK}{dt} = k_{KonK}[MatK][ProK] - k_{KoffK}[MatK\_ProK] - k_{KcatK}[MatK\_ProK] \quad (7)$$

$$\frac{dMatK\_MatK}{dt} = k_{KonK}[MatK]^2 - k_{KoffK}[MatK\_MatK] - k_{KcatK}[MatK\_MatK] \quad (8)$$

Independent Variables:

Scells = Number of cathepsin S producing cells (0, 1 or 2)

Kcells = Number of cathepsin K producing cells (0, 1 or 2)

Dependent Variables:

ProS = Free procathepsin S

MatS = Free mature cathepsin S

ProK = Free procathepsin K

MatK = Free mature cathepsin K

MatS_ProK = Mature cathepsin S bound to procathepsin K

MatS_MatK = Mature cathepsin S bound to mature cathepsin K

MatK_ProK = Mature cathepsin K bound to procathepsin K

MatK_MatK = Mature cathepsin K bound to mature cathepsin K

Parameters:

k_{Ssec} = rate of secretion of procathepsin S

k_{Sact} = rate of activation of procathepsin S to mature cathepsin S

k_{Sdeg} = rate of bulk loss of mature cathepsin S from the system

k_{Ksec} = rate of secretion of procathepsin K

k_{Kact} = rate of activation of procathepsin K to mature cathepsin K

k_{Kdeg} = rate of bulk loss of mature cathepsin K from the system

k_{SonK} = rate of binding of mature cathepsin S to pro or mature cathepsin K

k_{SoffK} = rate of unbinding of mature cathepsin S from pro or mature cathepsin K

k_{ScatK} = rate of catalysis of pro or mature cathepsin K by mature cathepsin S

k_{KonK} = rate of binding of mature cathepsin K to pro or mature cathepsin K

k_{KoffK} = rate of unbinding of mature cathepsin K from pro or mature cathepsin K

k_{KcatK} = rate of catalysis of pro or mature cathepsin K by mature cathepsin K

Model equations were implemented in COPASI [106]. Parameter estimation was carried out using COPASI's built-in genetic algorithm. The full details of parameter estimation methodology are described in Appendix 2.

4.2.7 Coculture of cathepsin overexpressing cells and macrophages with cathepsin inhibitor treatment and mRNA transfection

Thp-1 human monocytes were cultured in RPMI (Lonza, supplemented with 10% FBS, 1% L-glutamine, 1% Penicillin/Streptomycin) and differentiated into macrophages on 6 well plates at a density of 50,000 cells per well using 0.1uM phorbol 12-myristate 13-acetate (PMA) for 24 hours. Macrophages were switched to DMEM (Lonza, supplemented

with 10% FBS, 1% L-glutamine, 1% Penicillin/Streptomycin) for 24 hours, then wildtype or cathepsin K overexpressing HEK293T cells were added to the wells containing macrophages at a density of 50,000 cells and cultured overnight. The cells were washed with PBS and DMEM without FBS was added and collected after 24 hours. Cells were treated with vehicle, 50 μM E64 or cystatin C mRNA. Media was concentrated by adding 5 mL of cold acetone per 1 mL of cell culture media, incubating on ice for 30 minutes, then collecting the protein precipitate by centrifugation (3000g, 10 minutes, 4°C) and resuspending in a volume of lysis buffer resulting in a 10 fold concentration. Cells were lysed with zymography lysis buffer and sonicated. Concentrated media and lysate samples were stored at -80°C.

4.2.8 Transfection of cells with cystatin C mRNA

HEK293T cell Thp-1 macrophage cocultures were transfected with 8 μl of Lipofectamine 2000 was diluted in 100 μl of Opti-MEM reduced serum media and 4 μg of mRNA was diluted in 100 μl of Opti-MEM. Lipofectamine 2000 and diluted mRNA were mixed in equal amounts then Lipofectamine 2000-mRNA complexes were allowed to form for 5 minutes at room temperature. After incubation, 200 μl of complexes were added to each well for transfection. The total volume of media (DMEM and Opti-MEM) in transfected wells was equivalent to vehicle and E64 treated cells. For additional details on mRNA transfection, see Chapter 5.

4.2.9 Statistical analysis

Western blot band intensities were quantified by densitometry in ImageJ and data were normalized by dividing each band intensity on a gel by the intensity measured from either wildtype and cathepsin K or wildtype and cathepsin S HEK cells (setting these

groups to 1 across all membranes). Two-tailed Student's t-test with two-sample equal variance was performed on all statistical analysis.

4.3 Results

4.3.1 Cathepsins S and K interact in media in the presence of substrate

To test whether cathepsin S on K cannibalism can occur under cell culture conditions, recombinant cathepsins S and K were incubated in media alone and in combination with the fluorescently quenched substrate DQ Elastin and fluorescence was measured over time (Figure 4-1). Both cathepsins K and S degraded DQ Elastin, though cathepsin S produced much more fluorescence over a longer period than cathepsin K, consistent with known higher stability and activity of cathepsin S at neutral pH [29]. When cathepsins K and S were coincubated, greater substrate degradation was achieved than by either separately. However, the total fluorescence was less than predicted by adding together the fluorescence generated by each enzyme individually, indicating the two enzymes are interacting in some way. The predicted and actual cathepsin K and S outputs are quite similar for the first 30 minutes of incubation, but diverge by 40 minutes with the predicted K and S coculture fluorescence surpassing the actual K and S coculture measurements.

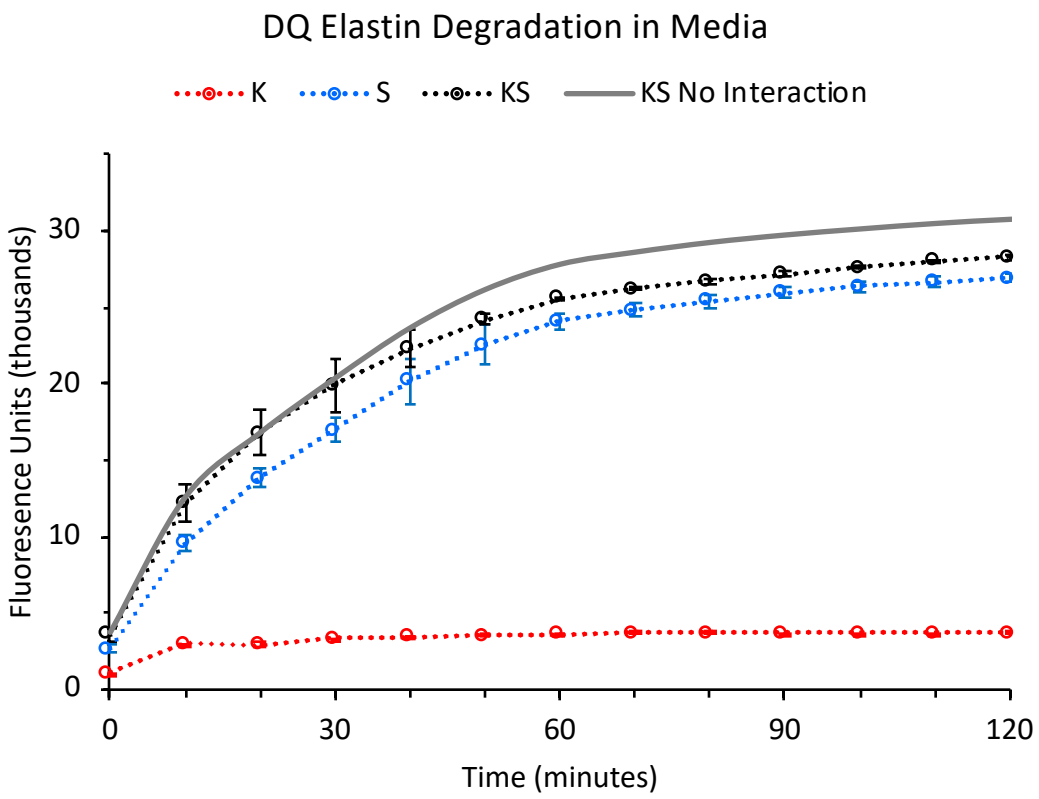

Figure 4-1: Cathepsins S and K interact in media in the presence of substrate.
Recombinant cathepsins K and S were incubated in cell culture media containing DQ Elastin and fluorescence was measured over time. Cathepsin S alone (blue dotted line) generated significantly more fluorescence than cathepsin K alone (red dotted line). When the two enzymes were incubated together they generated more fluorescence than either enzyme alone (black dotted line), however they generated less fluorescence than the expected sum of the two enzymes (gray solid line). The predicted fluorescence diverged from the measured coincubation after 30 minutes (n=2).

4.3.2 Coculture with cathepsin K and cathepsin S producing cells reduces procathepsin K in conditioned media

Cathepsin S and K cells, which were transfected to stably produce cathepsins under a constitutive promoter, were used as an experimental model of cathepsin cannibalism between tumor associated macrophages and breast cancer cells. Macrophages are known to secrete cathepsin S, particularly under inflammatory conditions, resulting in significant

pericellular elastinolysis [101]. Cathepsin K has been identified in primary breast tumors as well as metastatic sites in the bone [107]. Equal numbers of cathepsin S and K overexpressing HEKs were cocultured. Additional experimental groups substituted either CatS, CatK or both for wildtype (wt), CatS or CatK overexpressing cells resulting in six experimental groups: wt+wt, wt+CatS, wt+CatK, CatS+CatK, CatS+CatS, and CatK+CatK. We hypothesized that cathepsin S produced by cathepsin S cells would degrade cathepsin K produced by CatK cells, resulting in lower cathepsin K levels in media and lysate in the CatS+CatK group compared to the wt+CatK group.

Cells were seeded and cultured overnight, then switched to serum-free media for 24 hours to allow for protein secretion before media was collected to determine accumulation of cathepsins K and S in the media. Media was centrifugally concentrated and cells were lysed to determine if intracellular amounts of cathepsins K and S were affected by coculture. Protein concentration was measured by BCA, and equal amounts of lysate and media proteins were analyzed by Western blot (Figure 4-2). Procathepsin K was significantly lower in media when equal numbers of cathepsin K and S overexpressing cells were cocultured compared to equal numbers of cathepsin K overexpressing and wildtype cells, suggesting cathepsin S was able to degrade procathepsin K under normal cell culture conditions. This confirms the original hypothesis that cathepsin S is capable of cannibalizing cathepsin K in a cellular system. Mature cathepsin K, as well as pro- and mature cathepsin S, in the media was not significantly different between CatS+CatK and wt+CatK cells, suggesting cannibalism does not have a large impact on these species in this system, possibly due to the relatively low detection levels of mature cathepsin K in this system compared to procathepsin K.

Media levels of mature cathepsin K, and pro- and mature cathepsin S were highest in the experimental groups consisting of exclusively cathepsin K or cathepsin S overexpressing cells, respectively. Unexpectedly, media levels of procathepsin K were not significantly increased in the all cathepsin K cell group compared to the half cathepsin K and half wildtype group. The lack of an increase in procathepsin K upon doubling the CatK producing cells points to a nonlinear relationship between cathepsin K production and its accumulation in media. One potential explanation for this unusual relationship between K cells and CatK accumulation is mature cathepsin K cannibalizing procathepsin K, similarly to cannibalism by mature cathepsin S.

No significant differences in pro- or mature S and K were observed between lysate of CatS+CatK and wt+CatK cells, suggesting cathepsins were not endocytosed, or that endocytosed cathepsins do not significantly impact intracellular cathepsin levels in these cells. HEK293T cells are not known to be exceptionally endocytic and intracellular cannibalism caused by endocytosis is more likely to occur in highly endocytic cells such as macrophages. Intracellular pro and mature cathepsins K and S all increased when wildtype cells were replaced with cells producing the same cathepsin (CatK+CatK and CatS+CatS), suggesting any possible cathepsin K autodigestion has a minimal intracellular impact in these engineered cells.

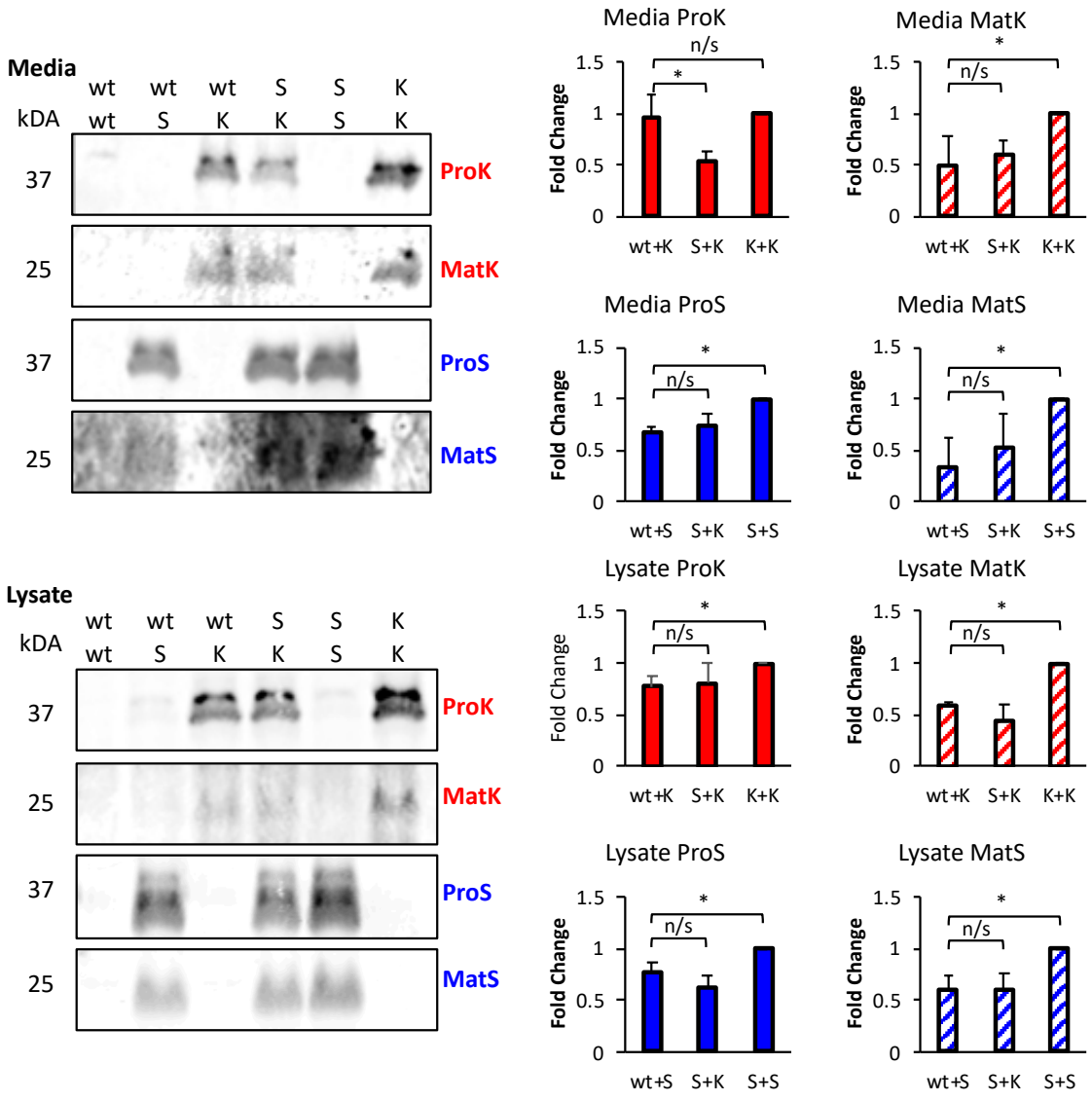

Figure 4-2: Media procathepsin K is significantly decreased when cocultured with cathepsin S producing cells. HEK293T cells were stably DNA transfected with plasmids to constitutively overexpress cathepsins K or S. Equal numbers of different combinations of cells were seeded, then switched to serum-free media for 24 hours. Lysate and conditioned media were collected and media was concentrated 10-fold using acetone precipitation then equal amounts of protein were analyzed by Western blot. Procathepsin K was significantly lower in media collected from cathepsin S and K cocultured cells (CatS+CatK) compared to K cells cocultured with wildtype cells (wt+CatK). Procathepsin K was not significantly increased when K cells filled the entire plate (CatK+CatK) compared to when half the plate was seeded with wildtype cells (CatK+wt). mature K, procathepsin S and mature S in the media and procathepsin K, mature K, procathepsin S and mature S had no significant difference between cathepsin K and S coculture compared to coculture with wildtype cells, but enzyme was significantly increased when wildtype

cells were replaced with more cathepsin overexpressors of the same type (n=3, * $p < 0.05$ by two tailed t-test).

4.3.3 Development of a mathematical model of cathepsin S and K dynamics in cellular coculture system

A mathematical model of cathepsin S and K overexpressing cells in coculture was constructed to explore possible cannibalistic interactions. Model interactions were implemented based on the diagram in Figure 4-3 and the mathematical representation is detailed in Methods. Cathepsin S and K cells are represented by independent variables, allowing the model to simulate the different coculture scenarios in Figure 4-2. Procathepsins S and K are secreted linearly, influenced by the amount of CatS or CatK cells in the system. Procathepsins are then activated to mature forms and degraded linearly. Mature cathepsins S and K are able to bind and degrade pro- and mature cathepsin K by forming complexes and degrading the enzymes according to enzyme-specific kinetic rates.

The model does not explicitly include cell growth, assuming rates of cell growth between S and K cells are similar over a 24-hour culture period. We assume cathepsins S kinetic rates (binding, unbinding and catalysis) on pro- and mature cathepsin K are equivalent, since the likely binding sites for cathepsin S on cathepsin K identified previously are unlikely to be affected by procathepsin K's propeptide [108]. Similarly, we assume the rates for cathepsin K binding and degrading procathepsin K are equivalent to the rates for cathepsin K binding and degrading mature cathepsin K.

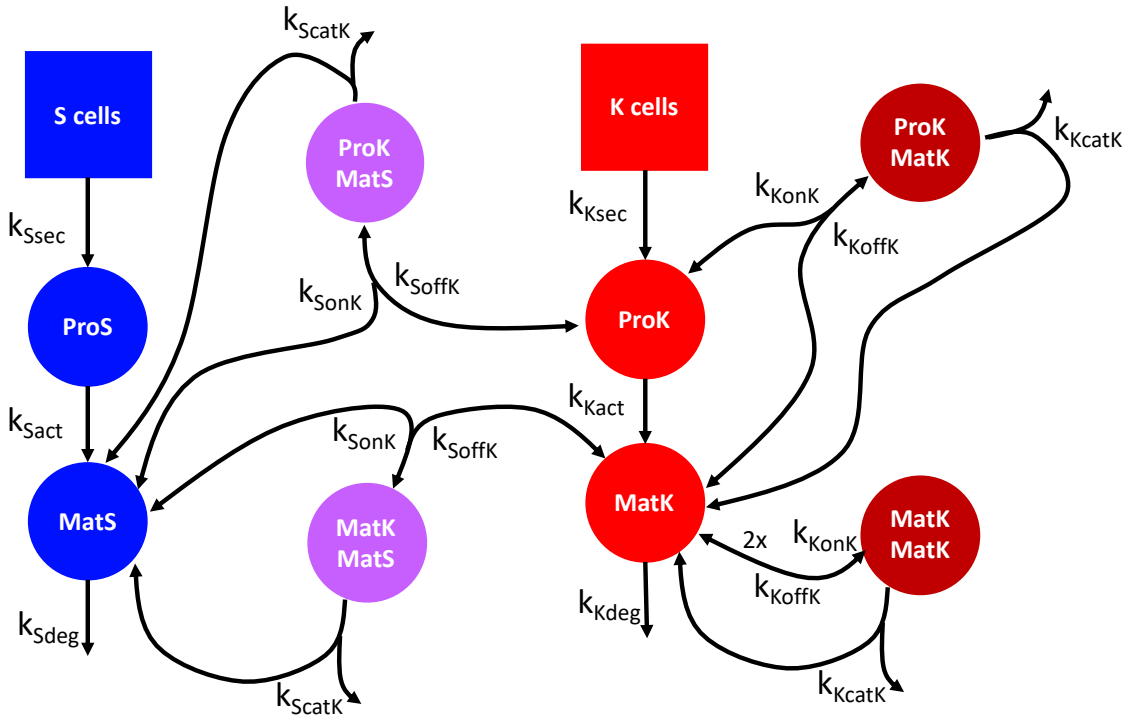

Figure 4-3: Model diagram for cathepsin S and K producing cells in coculture. S and K cells are modeled as independent variables secreting procathepsin S (ProS) and procathepsin K (ProK), which are activated to mature S (MatS) and mature K (MatK). MatS is able to bind and degrade both ProK and MatK. MatK is also able to bind ProK and MatK and degrade them as well.

4.3.4 Repeated parameter estimation yields four distinct families of solutions representing different degradation scenarios

The model was fit to a combination of the 24-hour data and time courses measuring media cathepsin accumulation by cathepsin overexpressing cells as described in Methods. To investigate the relative contributions of cathepsins S and K to cathepsin degradation, the 12 model parameters were estimated 100 times using COPASI's built-in genetic algorithm with a population size of 10,000 for 200 generations, requiring 2 million function evaluations per run (See Appendix 2 for parameter estimation details). Some parameters, including cathepsin S secretion, activation and degradation and cathepsin K secretion,

displayed little variation across the estimated solutions (Figure 4-4A). However, there was a much wider range of estimated values for the enzyme kinetic rates governing the binding and catalysis of cathepsin S on K cannibalism and cathepsin K autodegradation.

The estimated parameter sets were organized by hierarchical distance clustering, revealing 4 distinct families of solutions, red, cyan, green and blue (Figure 4-4B). A log scale heatmap of the parameter sets reveals that cathepsin K and S secretion rates are extremely similar across all solutions, along with CatS activation and CatS degradation (Figure 4-4B). Cathepsin K activation rate is consistently lower in the blue cluster compared to the red and green clusters indicating that the activation rate could be a pivotal difference in the behavior of these two clusters. Interestingly, cathepsin K degradation (not caused by K or S cannibalizing K) is extremely consistent among the red and green clusters, while it is much more varied among the blue cluster. Consistently high CatK activation and degradation rates in the green and red clusters suggests cathepsin K is being rapidly activated and degraded, likely without being proteolyzed by itself or cathepsin S.

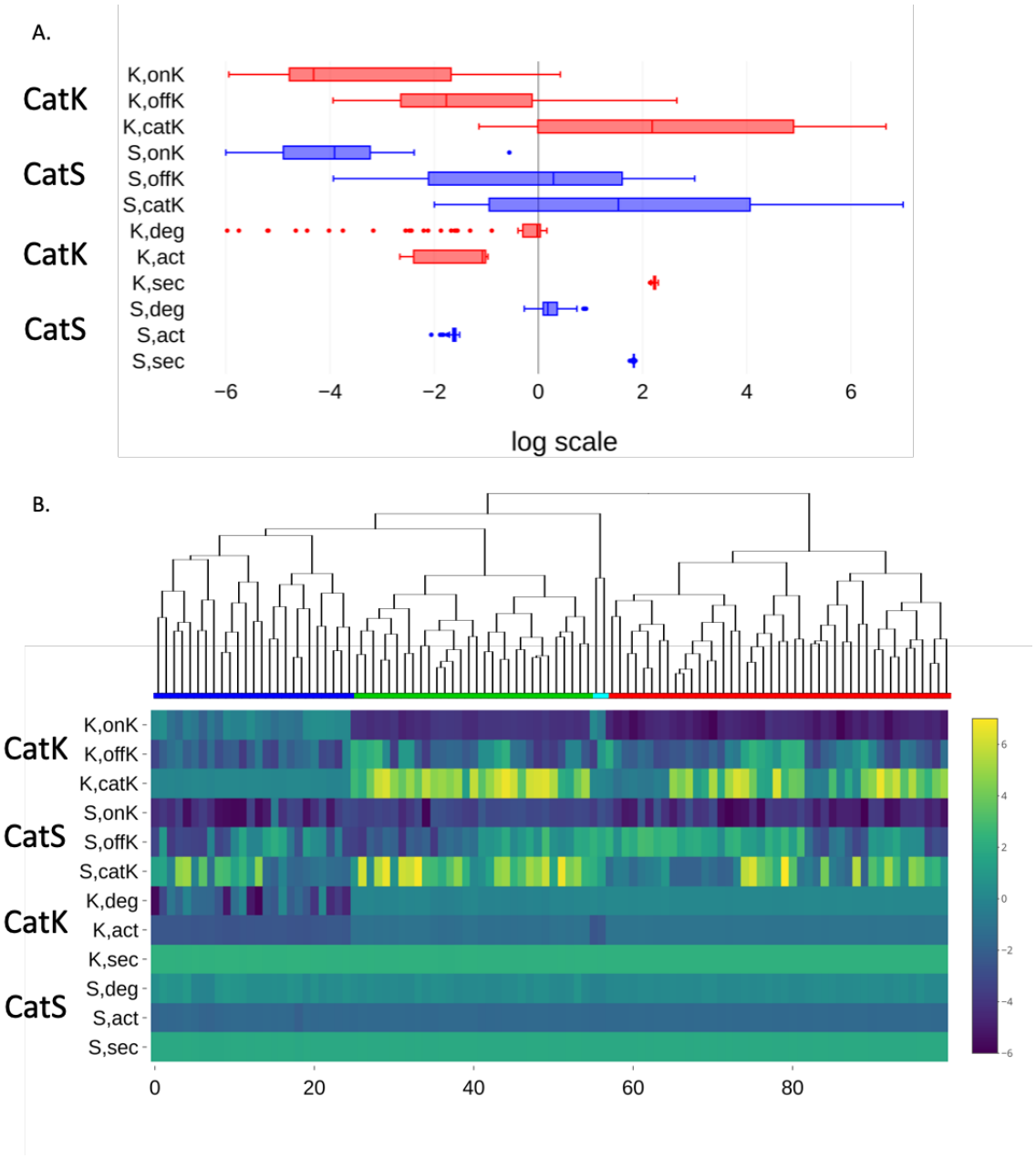

Figure 4-4: Repeated parameter estimation yields four distinct families of solutions. Parameters were estimated by genetic algorithm 100 times and the resulting parameter values were graphed by box plot (A). Estimated parameters were compared by hierarchical clustering and divided into four solution families and the values of parameters in those families were examined by heat map (B).

Simulations using the estimated parameters confirmed that the red and green solutions predicted much more cathepsin K activation and non-cathepsin mediated

cathepsin K degradation than the blue or cyan solutions (Figure 4-5A). Predicting high non-cathepsin mediated cathepsin K degradation is at odds with the experimental results shown in Figure 4-2, where adding cathepsin S producing cells significantly decreased media procathepsin K. The families of solutions were also analyzed based on the catalytic efficiency of cathepsin K degradation they predicted by cathepsins S and K. Kinetic parameters were calculated for each solution where:

$$K_m = \frac{(k_{off}+k_{cat})}{k_{on}} \qquad \text{and} \qquad \text{Catalytic Efficiency} = \frac{k_{cat}}{K_m}$$

Solutions were graphed according to predicted catalytic efficiency of cathepsin S on cathepsin K cannibalism against cathepsin K autodegradation (Figure 4-5B). Cathepsin K catalytic efficiency was considerably higher in the blue and cyan solutions, meaning cathepsin K autodegradation is more rapid in these solutions compared to the red and green parameter sets. The contribution of cathepsin S to cathepsin K degradation was less clearly delineated, with some red and blue solutions predicting low or high catalytic efficiency, while green and cyan solutions predicted a greater contribution of cathepsin S. These estimated parameter sets contained a large number of solutions that predicted cathepsins were minimally involved in cathepsin degradation (red and green clusters). Analysis of the clustered parameter heatmap revealed these solutions were characterized by high cathepsin K activation rates, which resulted in the high cathepsin K activation and subsequent non-cathepsin mediated cathepsin K degradation we observed in simulations.

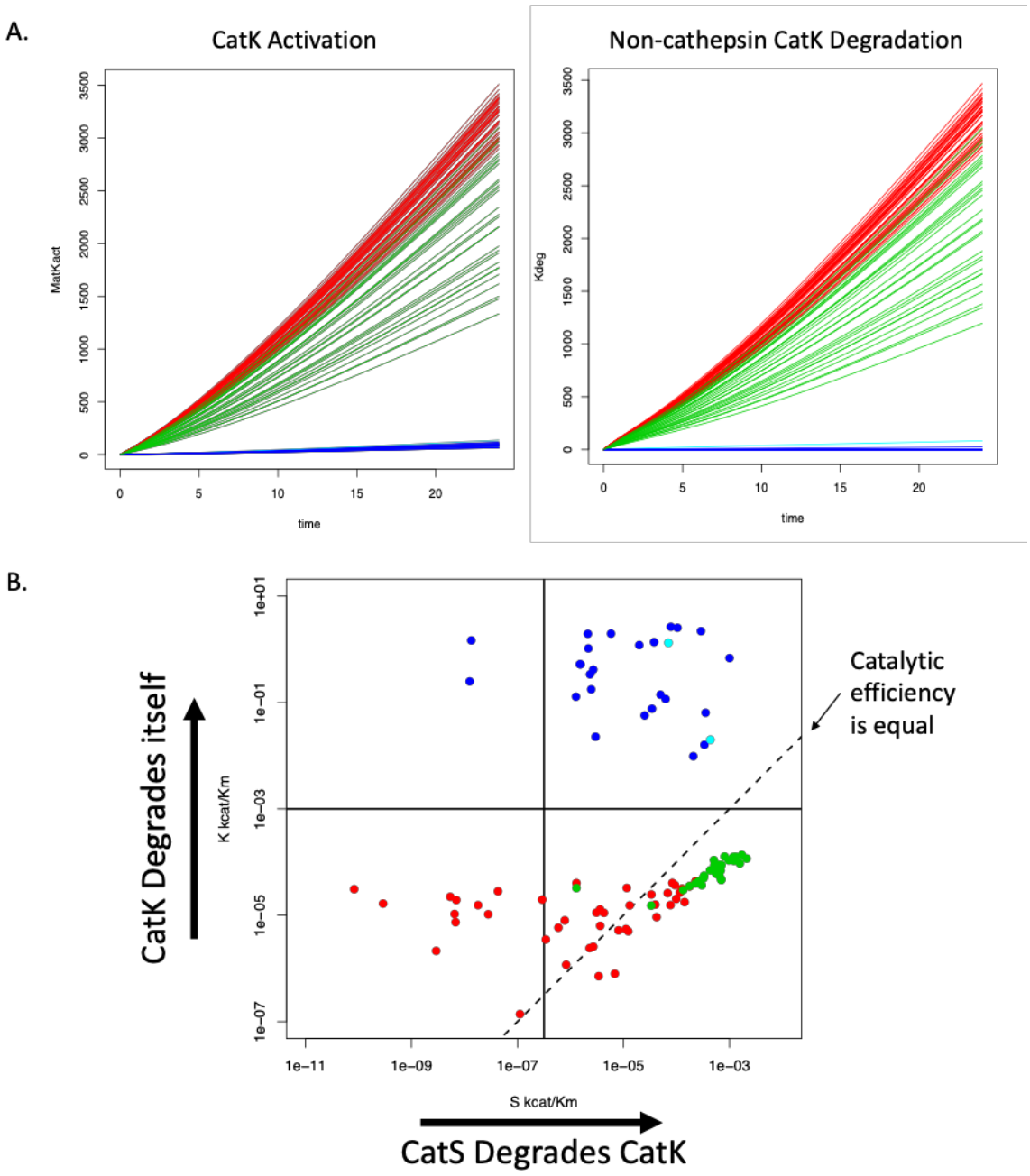

Figure 4-5: Parameter solution families predict different degrees of cathepsin involvement in cathepsin K degradation. Time course results were simulated for each family of solutions and the amount of cathepsin K activation, and non-cathepsin mediated cathepsin K degradation were graphed. The red and green solutions predicted much larger cathepsin K activation and non-cathepsin mediated degradation than the blue or cyan solutions (A). Parameter solution families predict different catalytic efficiency of cathepsin K on cathepsin K and cathepsin S on cathepsin K. The catalytic efficiency of each solution was plotted and red and green solutions were found to predict low cathepsin K on cathepsin K

catalytic efficiency, while blue and cyan solutions predicted high cathepsin K on cathepsin K catalytic efficiency (B).

4.3.5 Restricting parameter space through cathepsin K activation yields solutions which predict cathepsin involvement in cathepsin degradation

To better identify solutions capable of explaining the relative contributions of cathepsins S and K to cathepsin K cannibalism, we refined our parameter estimation bounds by lowering the upper limit of the cathepsin K activation rate and repeated the parameter estimation. Analysis of the catalytic efficiencies of the new parameter sets revealed three main solution families, CatK dominant, equal partners and CatS featured (Figure 4-6). Notably, all of the refined parameter solutions predicted cathepsin degradation was influenced by cathepsins, agreeing with experimental observations. However, these solutions differ in their predicted contributions of cathepsins S and K to cathepsin K degradation.

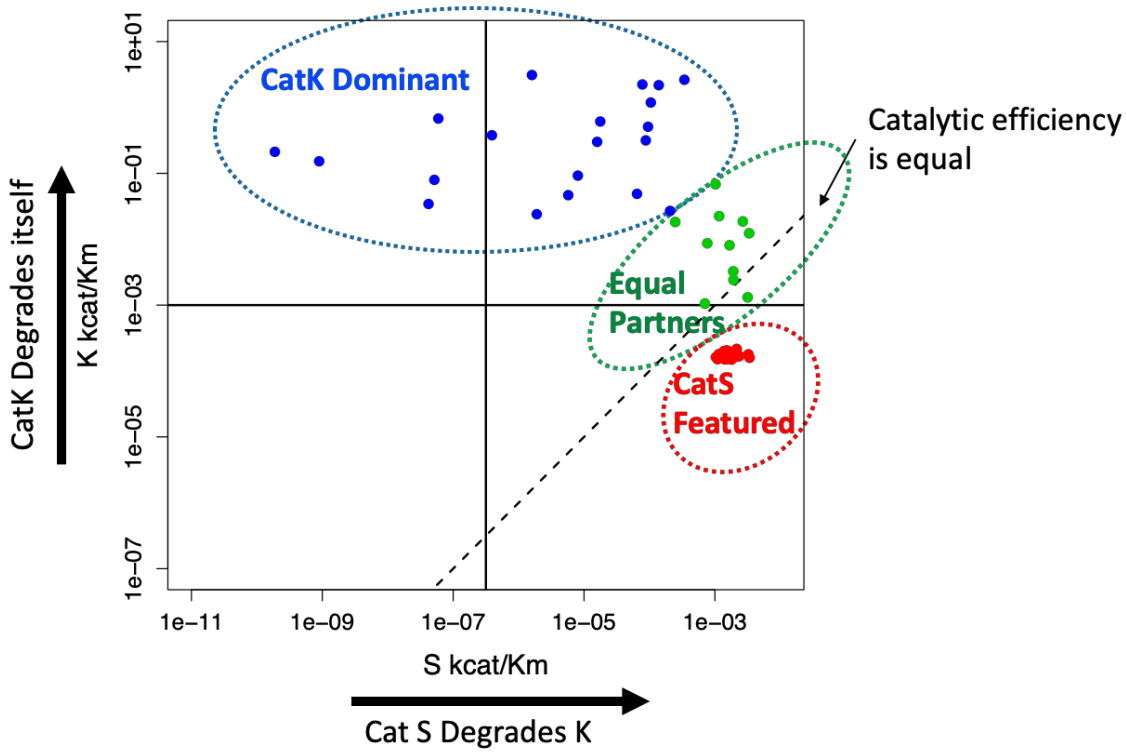

Figure 4-6: Restricting parameter estimation space yields three new solution families. Parameter estimation was repeated after lowering the bounds on cathepsin K activation rate, which was found to be high among solutions predicting little cathepsin involvement in cathepsin degradation. The refined parameter estimation yielded CatK Dominant, Equal Partners and CatS Featured solutions, which could be differentiated based on catalytic efficiency of cathepsins S and K on degrading cathepsin K.

Each solution was simulated to show predicted procathepsin K degradation when cathepsin S and cathepsin K overexpressing cells are cocultured. The CatS featured solutions (red) predict the least degradation of procathepsin K by cathepsin K, and predict roughly the same degradation of procathepsin K by cathepsins K and S (Figure 4-7A). The CatS featured solutions are also quite internally consistent, with most of the simulations among this group predicting very similar behavior. In contrast, the CatK dominant and equal partners solutions predict a wider range of possible procathepsin K degradation, but

in general they predict more involvement of cathepsin K than cathepsin S in procathepsin K degradation.

All three solution families fit the 24-hour time course data similarly well, but the solutions predicted drastically different behavior for time points outside the experimental time course (Figure 4-7B). CatS featured predictions were again quite internally consistent and display far less volatility and oscillatory behavior than CatK dominant or equal partners solutions. Interestingly, the CatS featured solutions reach a reasonable steady state even when cathepsin S producing cells are not present in the system. When cathepsin S cells are added to the system, the amount of procathepsin K continues to decrease over time as the total cathepsin S accumulates in the media. While oscillatory behavior is possible in the system, the extreme rises and falls predicted by the CatK dominant solutions are unlikely to occur biologically.

Despite the consistency in the CatS featured predictions and the similar catalytic efficiencies predicted in this group shown in Figure 4-6, there are a range of enzyme kinetic parameters observed among these solutions. A box plot of all of the refined solutions (CatK dominant, equal partners and CatS featured) shows there is still considerable variation among solutions, particularly in enzyme kinetic rates (Figure 4-8A). CatS featured solutions restrict these ranges somewhat, particularly among the on rates for cathepsins S and K binding to cathepsin K (Figure 4-8B). However, off and catalytic rates retain a great deal of variation even among the CatS featured solutions.

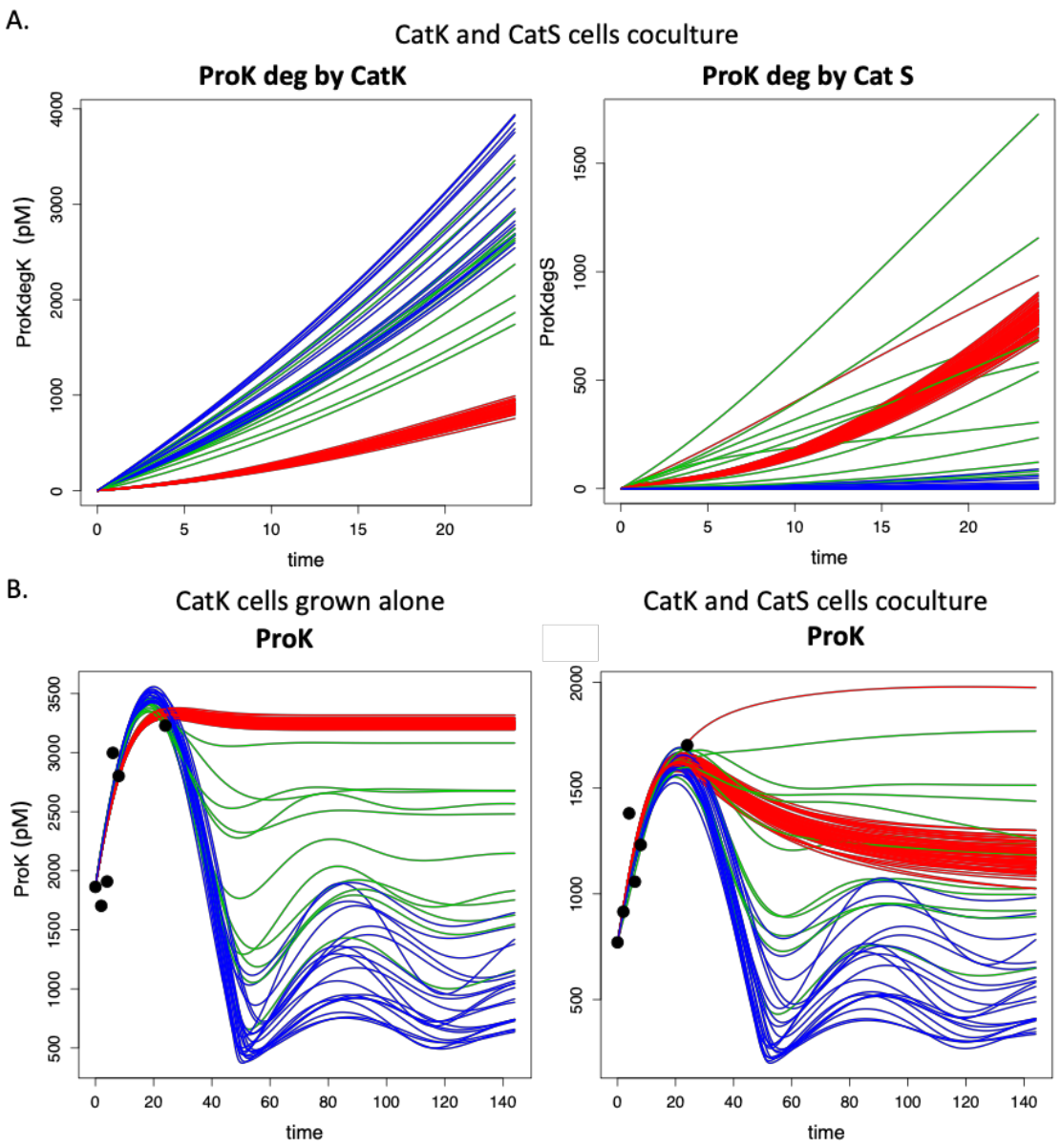

Figure 4-7: CatS featured parameters are more consistent and less volatile than other solutions.

Model time course simulations were generated for each of the refined solutions. CatS featured solutions (red) predicted lower contributions of cathepsin K, and higher contributions of cathepsin S, to procathepsin K degradation compared to the CatK dominant (blue) and equal partners (green) solutions (A). All solutions fit the time course data similarly over 24 hours, but different parameter sets predicted very different dynamic behavior once the 24 hour time course ended (B). Most CatS featured solutions were tightly grouped, changing gradually at the end of the time course. The CatK dominant and equal partners solutions displayed more variance and volatility at the end of the time course.

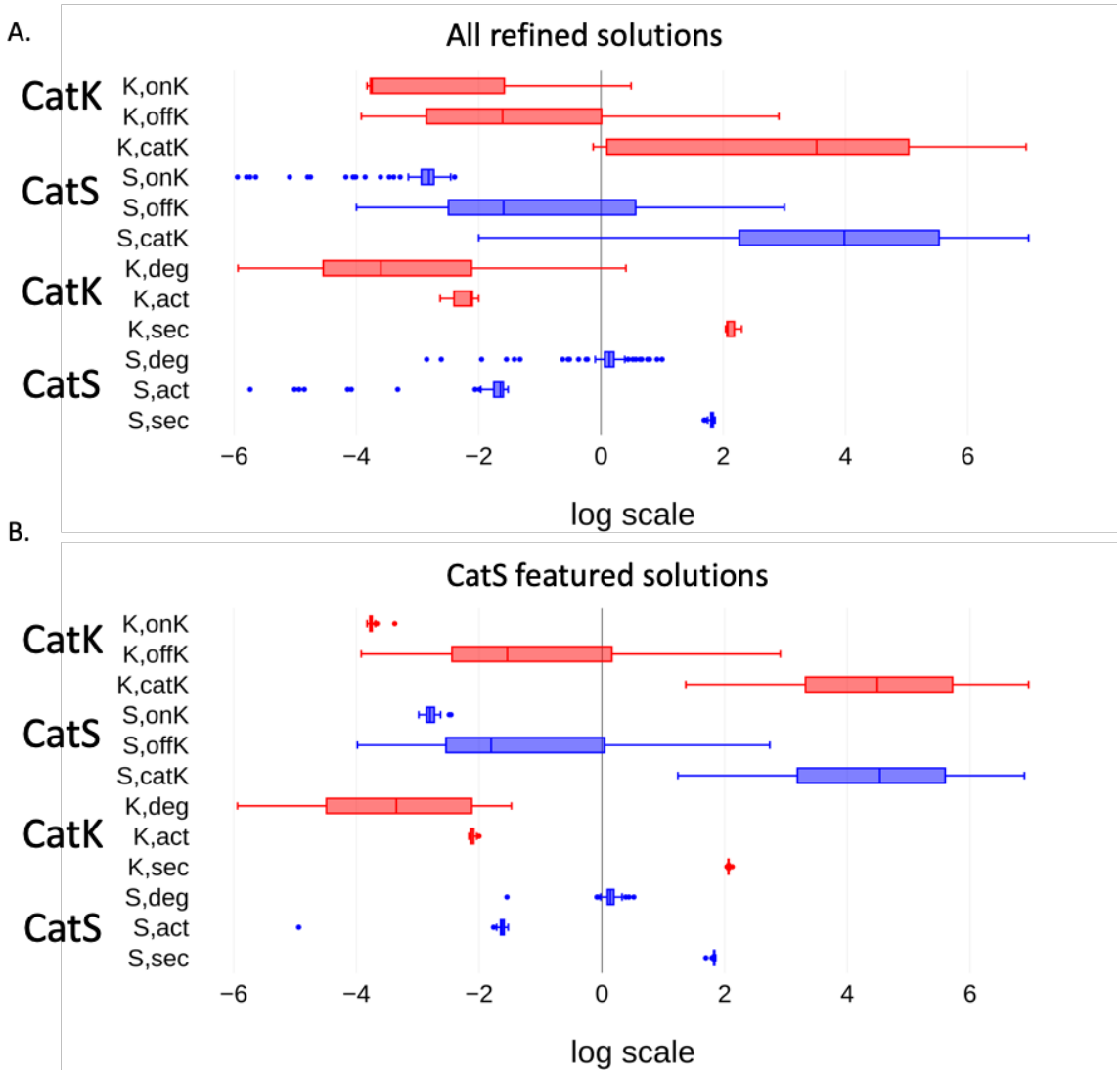

Figure 4-8: CatS featured solutions have consistent kinetic on rates.
The values of parameters among the refined solutions (A) and CatS featured solutions (B) were examined by box plot. CatS featured solutions displayed less variance than the combination of all refined solutions, most notably among on rates for cathepsins K and S binding to cathepsin S.

4.3.6 Incubation with cathepsin inhibitors increases procathepsin K accumulation in conditioned media of macrophages cocultured with cathepsin overexpressing cells

To test whether cathepsins produced by macrophages are capable of cannibalizing cathepsin K, Thp-1 human macrophages were cocultured with cathepsin K overexpressing

cells. Cells were treated with the cathepsin inhibitor E64 and transfected with mRNA for the endogenous protein inhibitor of cathepsins, cystatin C, which is constitutively secreted by most human cells. Treating cells with E64 was found to increase media procathepsin K compared to vehicle treatment, consistent with E64 preventing cathepsin S and cathepsin K degradation of procathepsin K (Figure 4-9). Cystatin C mRNA did not affect level of procathepsin K in the conditioned media, but it did significantly decrease the level of procathepsin S detected in conditioned media (Figure 4-9B). Cystatin C mRNA was confirmed to cause increased amounts of cystatin C in the media compared to vehicle treated cells.

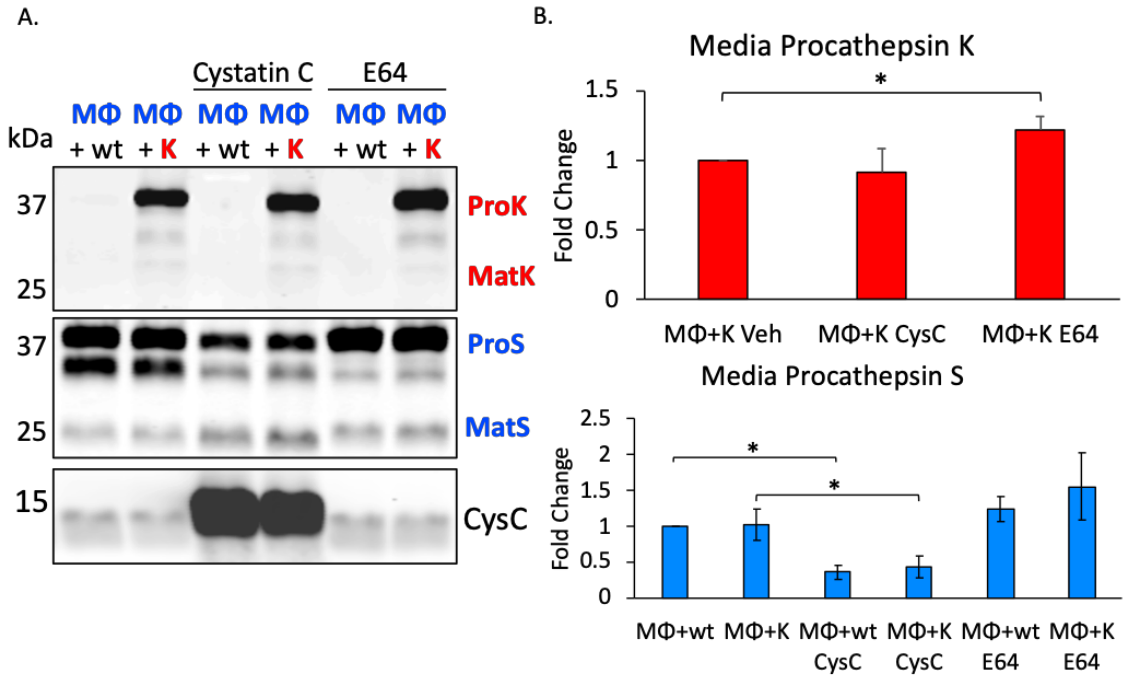

Figure 4-9: E64 prevents degradation of media procathepsin K when macrophages are cocultured with K+ cells.
Thp-1 macrophages were cocultured with wildtype or cathepsin K producing HEK cells, then treated with cystatin C mRNA or E64. Conditioned media were collected after 24 hours and cathepsin K, cathepsin S and cystatin C were measured by Western blot (A) and quantified by densitometry (B). Procathepsin K was significantly increased following E64 treatment, while cystatin C mRNA had no effect on cathepsin K levels. Cystatin C significantly decreased procathepsin S in media from wildtype and cathepsin K overexpressing cells, and E64 had no effect on cathepsin S levels. Probing for cystatin C

confirmed cystatin C mRNA transfection induces cystatin C overexpression (n = 3, * p < 0.05 by t-test).

Lysate from the macrophage-HEK293T cocultures was probed for cathepsins K and S and cystatin C to evaluate the effects of inhibitor treatment and mRNA transfection on intracellular protein levels (Figure 4-10). Cathepsin K did not appear to change under any treatment conditions. Interestingly, pro- and mature cathepsin S was almost undetectable following cystatin C transfection but appeared unaffected by E64 treatment. Cystatin C mRNA was again confirmed to increase cystatin C expression, where the protein was only detectable in lysate from mRNA transfected cells. Additional experiments using different concentration methods and carboxymethyl papain to compete with cystatin C, were conducted and shown in Appendix 5.

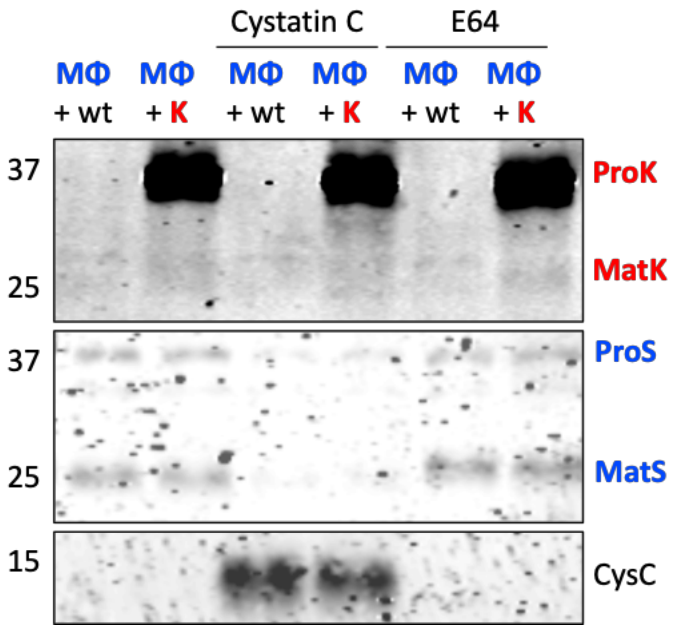

Figure 4-10: Transfection with cystatin C mRNA decreases cathepsin S in macrophage lysate. Thp-1 macrophages were cocultured with wildtype or cathepsin K producing HEK cells, then treated with cystatin C mRNA or E64. Cells were lysed after 24 hours and cathepsin K, cathepsin S and cystatin C were measured by Western blot. Pro- and mature cathepsin S levels were decreased following cystatin C mRNA treatment, while E64 had no effect on cathepsins S or K in lysate. Cystatin C was detectable in the lysate of cystatin

C mRNA treated cells, while cystatin C was undetectable in cells not transfected with mRNA.

4.4 Discussion

We sought to investigate the effects of cathepsin cannibalism in multicellular systems where it regulates proteolysis through the controlled degradation of proteases. We showed that cathepsin cannibalism between recombinant cathepsins S and K likely occurs in cell culture media, and in an engineered cellular system. Using mathematical modeling techniques we identified ranges of likely kinetic rates for cathepsin cannibalism, predicting that in the extracellular environment, cathepsin S on cathepsin K cannibalism proceeds more rapidly than cathepsin K autodegradation. We also demonstrated that cathepsin cannibalism occurs in a macrophage coculture system and it can be prevented using cathepsin inhibitors. These results provide further evidence that cathepsin cannibalism and autodegradation are physiologic regulatory processes which can occur in the shared extracellular spaces of cells producing these proteases, and that this regulation could be disrupted by treatment with cathepsin inhibitors potentially causing unexpected side effects including increases in cathepsins.

We first established that cathepsin cannibalism between cathepsins S and K can occur in the presumed neutral pH environment of cell culture media. We found cathepsin S was considerably more active than cathepsin K under these conditions, and that when coincubated together cathepsins K and S hydrolyze less of the fluorogenic substrate than would be predicted by adding the outputs of the two enzymes together (Figure 4-1). Next, we created an experimental model of macrophage and breast cancer cell cathepsin dynamics using HEK293T cells engineered to overexpress cathepsins S and K. We found coincubation of cathepsin K and S producing cells decreased procathepsin K in the cell

culture media, supporting the existence of cathepsin cannibalism in cellular systems (Figure 4-2). We also observed doubling the concentration of cathepsin K producing cells did not significantly increase the procathepsin K concentration in the media, suggesting a nonlinear relationship between cathepsin K production and accumulation, hypothesized to be caused by autocatalysis of cathepsin K by itself.

We created a computational model of cathepsin S and K coculture to examine the differences in degradation of cathepsin K by itself compared to degradation by cathepsin S (Figure 4-3). We parameterized the model to time course data, generating four distinct families of solutions through hierarchical clustering, each predicting different contributions of cathepsin K and cathepsin S to the degradation of cathepsin K (Figure 4-4). Many of these initial solutions predicted cathepsins were not responsible for significant cathepsin K degradation, a prediction in conflict with the experimental evidence (Figure 4-5). We were able to refine our solutions by restricting the values of cathepsin K activation, resulting in a set of refined solutions (CatK dominant, equal partners and CatS featured) that predicted different contributions of cathepsins to cathepsin K degradation (Figure 4-6). CatS featured solutions generated the most consistent and least volatile predictions, providing evidence for their validity in explaining the biological system (Figure 4-7).

To demonstrate that cathepsin cannibalism can occur with the involvement of macrophages, Thp-1 macrophages were cocultured with cathepsin K overexpressing cells and cathepsin degradation was inhibited with E64 and cystatin C overexpression. E64 increased procathepsin K, consistent with inhibition of cathepsins preventing cathepsin cannibalism (Figure 4-9). However, cystatin C mRNA transfection did not affect cathepsin K in the media but did decrease cathepsin S in the media and lysate of cocultured cells

(Figure 4-10). This decrease in lysate cathepsin S following cystatin C mRNA transfection could be evidence for a regulatory mechanism controlling cathepsin S production modulated by cystatin C levels. Additional experiments measuring cathepsin S protein and mRNA levels, utilizing recombinant cystatin C and different dosages of cystatin C mRNA will need to be conducted to determine the cause of this surprising decrease in lysate cathepsin S following cystatin C mRNA transfection.

The primary goal of representing the cathepsin coculture system mathematically was parsing the relative contributions of cathepsins S and K to cathepsin K degradation. Through multiple rounds of parameter estimation and refinement, we were able to arrive at a family of solutions that predict cathepsin S degrades cathepsin K with somewhat more rapid kinetics than it degrades itself. We were also able to derive some other insights about the system through analyzing these parameters. Most of the solutions we generated predicted the secretion rate of cathepsin K as between 50-100% faster than cathepsin S secretion. This seems to likely be the case as media cathepsin K levels were generally higher than cathepsin S in spite of the hypothesized degradation of cathepsin K by itself and cathepsin S. Additionally, there was very little disagreement on secretion rates between the different parameter sets (Figure 4-4A and Figure 4-8A). Previous *in vitro* results suggest cathepsin K autodegradation could actually be faster than cathepsin S on K cannibalism (Ferrall-Fairbanks et al. in preparation). These *in vitro* experiments were conducted at pH 6, while cathepsin K is much less active in the slightly alkaline pH environment of cell culture media. However, a significant amount of enzyme degradation could be happening in the pericellular space, where the pH can be lower due to the secretion of acidic metabolites [109].

5 PERTURBATIONS TO CATHEPSIN PROTEOLYTIC NETWORKS WITH MOLECULAR APPROACHES

5.1 Introduction

DNA transfection and the generation of stable cell lines are incredibly powerful techniques that have allowed us to generate experimental models of cathepsin overexpression (Chapter 4) and purify recombinant cathepsins from conditioned media (Chapter 6). However, DNA transfection is limited by its efficiency, particularly in primary cells, which do not readily express foreign DNA. To address this need we collaborated with the Santangelo Lab at Georgia Tech to produce a series of cathepsin and cystatin C specific mRNA suitable to induce overexpression in a variety of cell types with high efficiency [110]. Additionally, these cathepsin and cystatin mRNA could be used to probe for possible regulatory connections between different cathepsin family members, where the overexpression of one cathepsin could decrease the expression of another through digestion or potentially increase expression through stimulating transcription or inhibiting degradation of another cathepsin.

5.2 Materials and Methods

5.2.1 Development of cathepsin and cystatin C mRNA for overexpression

Overexpression mRNA coding for human wildtype cathepsins K, L, S and V, cystatin C and mutated cathepsin K was developed using methodology as previously described [111].

5.2.2 Transfection of cells with cathepsin and cystatin C mRNA

HEK293T cells were grown to approximately 90% confluence on 12 well plates containing 1mL of DMEM (Lonza, supplemented with 10% FBS, 1% L-glutamine, 1% Penicillin/Streptomycin) before transfection. For each well, 4 µl of Lipofectamine 2000 was diluted in 50 µl of Opti-MEM reduced serum media and 2 µg of mRNA was diluted in 50 µl of Opti-MEM. Diluted Lipofectamine 2000 and diluted mRNA were mixed in equal amounts then Lipofectamine 2000-mRNA complexes were allowed to form for 5 minutes at room temperature. After incubation, 100 µl of complexes were added to each well for transfection. Cells were lysed after 12 hours and overexpression was assessed with Western blot and cathepsin zymography.

5.2.3 Western blot

Western blot was used to measure mRNA induced protein overexpression. Total protein concentration in cell lysate was determined by BCA, then equal amounts of protein were mixed with reducing loading dye, briefly boiled, then loaded into 12.5% polyacrylamide gels and separated by electrophoresis. Proteins were transferred onto nitrocellulose membranes, blocked and probed with primary antibodies for human cathepsin S (R&D Systems), human cathepsin L (R&D Systems), human cathepsin V(R&D Systems), human cathepsin K (Proteintech) or human cystatin C (EMD Millipore). Membranes were probed with secondary antibodies (LI-COR) then imaged on a LI-COR Odyssey CLx Imaging System.

5.2.4 Cathepsin zymography

Cathepsin zymography was used to measure active cathepsins in mRNA transfected cells as described previously [84, 85]. Briefly, equal amounts of protein were loaded in

12.5% polyacrylamide gels embedded with 0.2% gelatin and separated by electrophoresis at 4°C. Gels were washed in renaturing buffer and incubated in assay buffer overnight at 27°C. Gels were stained with Coomassie blue, destained and imaged on an ImageQuant LAS 4000 (GE Healthcare).

5.2.5 Measuring proteolysis with fluorescently quenched elastin

HEK293T cells were transfected with mRNA then after 4 hours cells were switched to media containing 0.05mg/ml DQ elastin. Samples of the media were taken after another 4 and 20 hours, and fluorescence was quantified in a plate reader. Background fluorescence (measured with a media containing DQ Elastin control) was subtracted out of all reported measurements.

5.3 Results

5.3.1 Transfection with cathepsin mRNA induces overexpression in cells

Cathepsin K, L, S, V and cystatin C mRNA capable of inducing overexpression were developed to study cathepsin cannibalism in living cells. Additional mRNA for cannibalism resistant mutants of cathepsin K (K^{V171A}, K^{L253V}, K^{L253A}) were developed according to predictive cleavage sites previously identified using the protease-on-protease predictive algorithm, PACMANS [108]. HEK293T cells were transfected with cathepsin and cystatin C mRNA and efficacy of overexpression was evaluated by Western blot (Figure 5-1). Procathepsins L, S and V and cystatin C were detectable in transfected lysate. Procathepsin K^{V171A} and K^{L253V} were also detectable by Western blot, however wildtype cathepsin K and K^{L253A} were not. Mature cathepsin S and V were clearly detectable, but no

other mature cathepsins were detectable in mRNA transfected HEK293T cells, indicating the mRNA protein is not being activated properly in these cells.

MDA-MB-231 breast cancer cells were also transfected with cathepsin and cystatin C mRNA to determine the effectiveness of mRNA transfection in a cell type that expresses multiple cathepsins endogenously. We also wanted to test the hypothesis that overexpression of cellular cathepsins could affect the expression of other cathepsins. MDA-MB-231 cells were transfected, lysed and protein expression was evaluated by Western blot (Figure 5-2A). Endogenous procathepsin K, as well as pro and mature cathepsins L and S were detectable in all MDA-MB-231lysate. Similarly to HEK293T cells, overexpression of procathepsins L, S and V, K^{V171A} and K^{L253V} were detected in transfected MDA-MB-231cells. Wildtype cathepsin K and K^{L253A} were potentially increased following transfection, but to a lesser extent than K^{V171A} and K^{L253V}. Mature cathepsin S was increased following transfection, but no other mature cathepsins appeared elevated compared to control and vehicle treated cells. Procathepsin V appeared to be detectable in cells transfected with K^{V171A} and K^{L253V} mRNA, however a subsequent Western blot using recombinant K and lysate from engineered HEK293T cells that stably express cathepsin K, revealed the cross reactivity of the cathepsin V antibody used previously (Figure 5-2B).

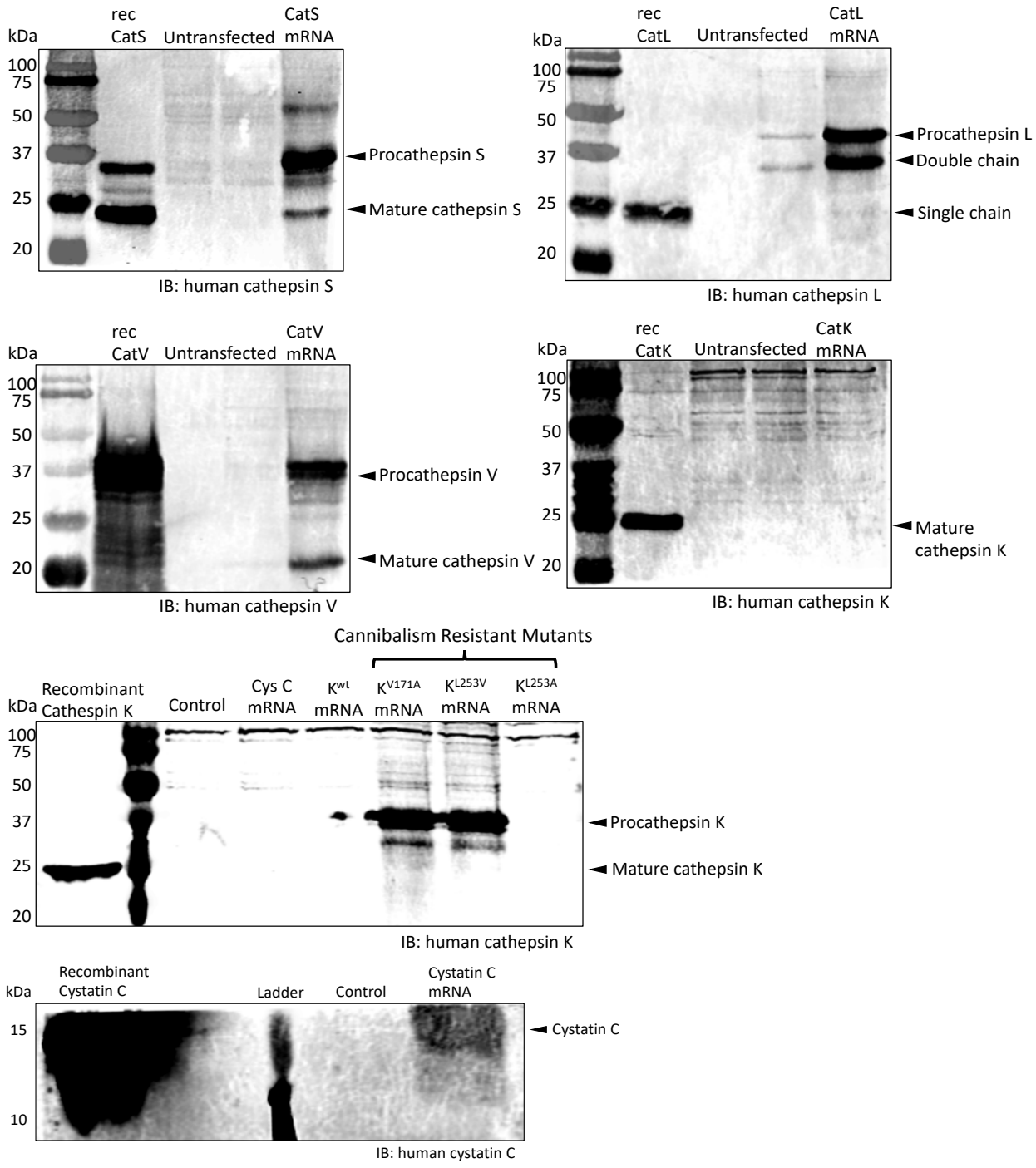

Figure 5-1: mRNA transfected cells produce cathepsins S, L, V, cannibalism resistant mutants of cathepsin K and cystatin C. HEK29T were transfected with mRNA for cathepsins S, L, V, K^{wt}, K^{V171A}, K^{L253V} K^{L253A} or cystatin C. After 12 hours, cells were lysed and equal amounts of protein were analyzed by Western blot. Transfected protein was clearly detectable in cells transfected with mRNA for cathepsin S, L, V, K^{V171A}, K^{L253V} and cystatin, but not cathepsin K^{wt} or K^{L253A}.

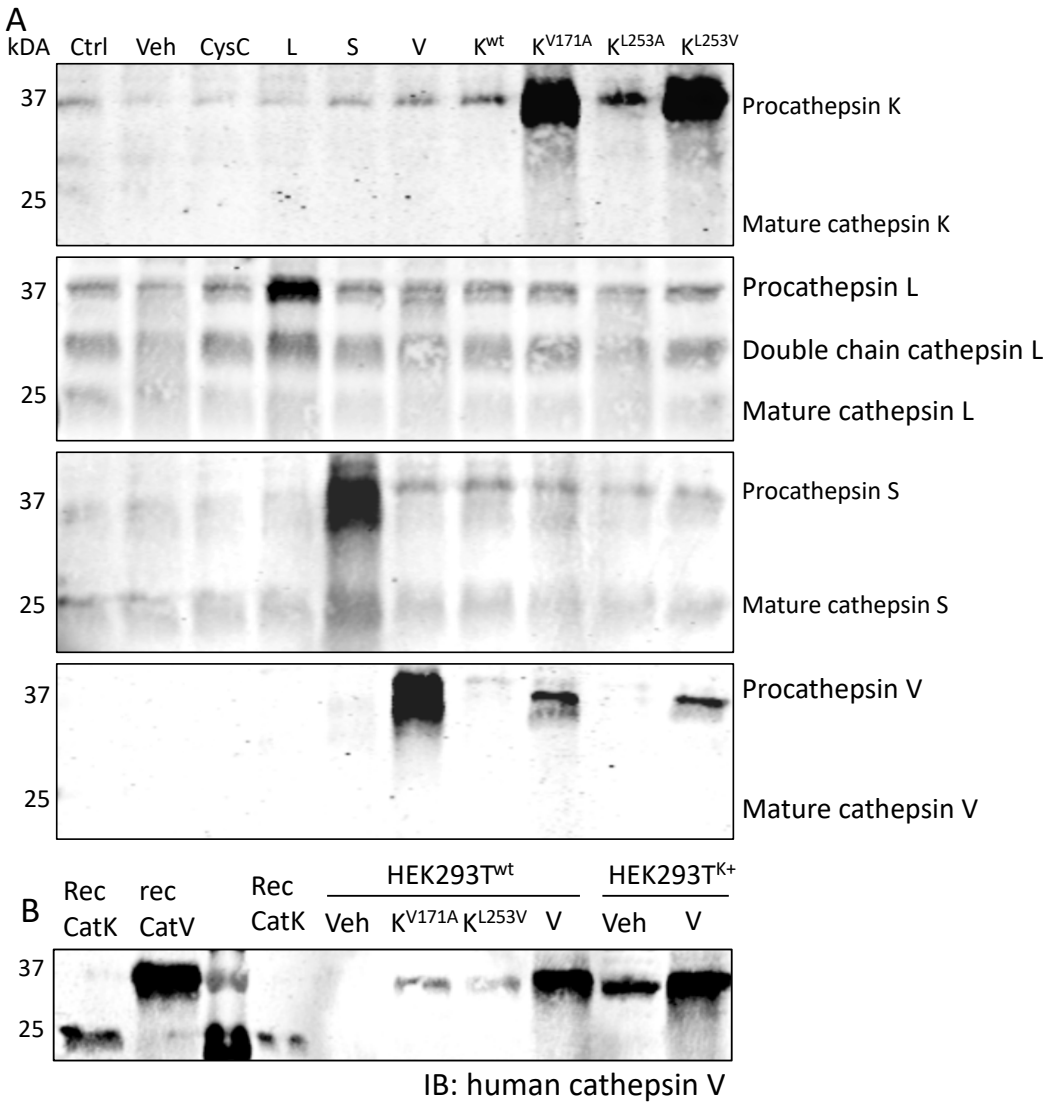

Figure 5-2: mRNA transfection is successful in breast cancer cells. MDA-MB-231 breast cancer cells were transfected with cathepsin overexpression mRNA, lysed and equal amounts of protein were measured by Western blot (A). Procathepsin K was elevated in K^{V171A} and K^{L253V}, but not K^{wt} or K^{L253A}, but no mature K was detectable in any lysate. Procathepsin L and procathepsin S were elevated in cells transfected with L and S mRNA, respectively. Procathepsin V appeared to be elevated in V, K^{V171A} and K^{L253V} mRNA transfected cells. Recombinant cathepsin K and lysate from $HEK293T^{K+}$ cells shows the cathepsin V antibody cross reacts with cathepsin K.

5.3.2 mRNA transfection yields active cathepsin V and inhibitory cystatin C

To test whether mRNA transfection can produce active cathepsins capable of degrading substrate, mRNA transfected cells were lysed 12 and 24 hours post transfection and active cathepsins were assayed using cathepsin zymography (Figure 5-3). An active cathepsin V band was detectable in cells transfected with cathepsin V mRNA. Active cathepsin L bands were detectable in vehicle and mRNA transfected cells, though the active cathepsin L band was not increased in the cells transfected with L mRNA. No cathepsin S or cathepsin K bands were detectable by zymography.

To test whether mRNA transfected cells can secrete active cathepsins and inhibitory cystatin C, HEK293T cells were transfected with mRNA for cathepsin S, cystatin C or both cathepsin S and cystatin C, then switched to media containing 50μg/mL DQ Elastin after 4 hours. Media was collected after another 4 and 20 hours, and fluorescence was immediately measured using a plate reader. Fluorescence was not increased in cells transfected with cathepsin S mRNA, suggesting transfected cells do not secrete catalytically active cathepsin S. However, cystatin C transfected cells had markedly decreased fluorescence compared to the other experimental groups after 20 hours, indicating cystatin C produced by transfected cells is capable of inhibiting proteolysis by secreted cellular proteases.

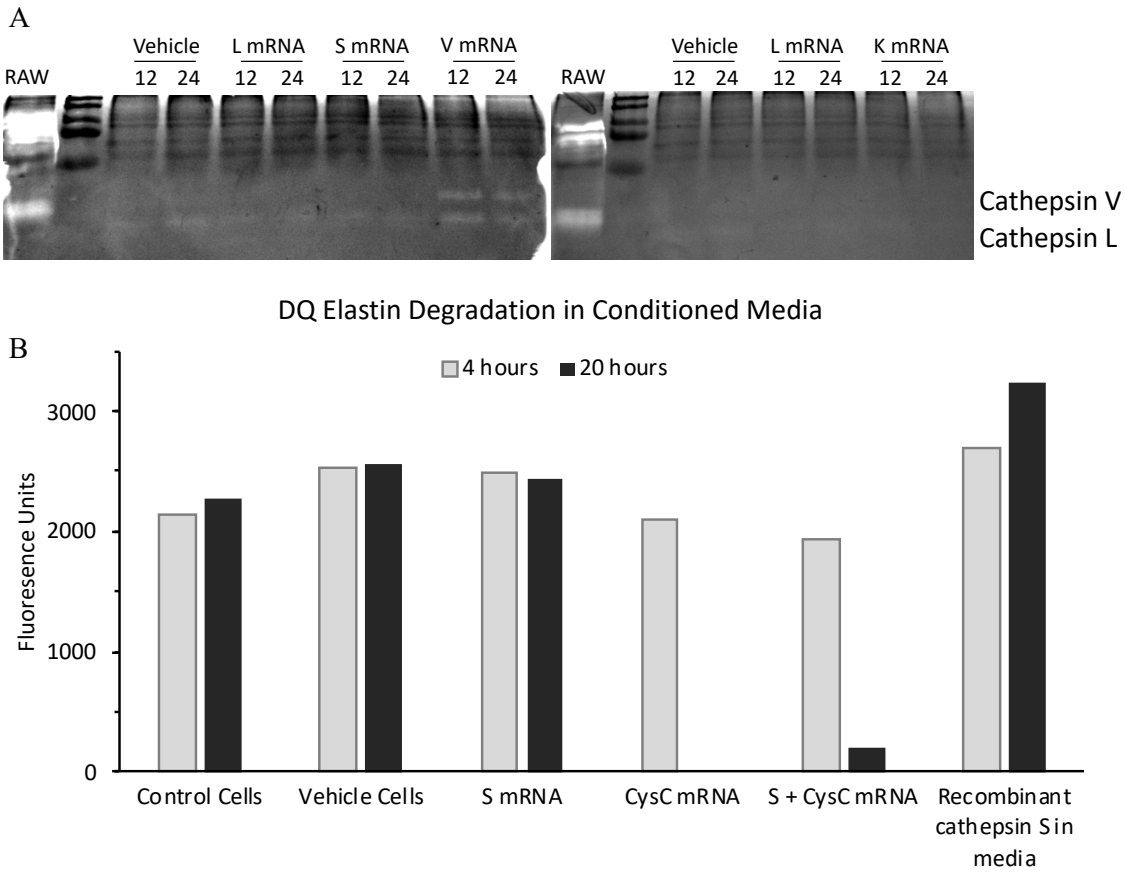

Figure 5-3: mRNA transfected cells produce active cathepsin V as well as cystatin C capable of inhibiting elastin degradation. HEK293T cells were transfected with mRNA for cathepsins K, L S and V, lysed and equal amounts of protein were loaded for cathepsin zymography. Cathepsin V transfected cells produced an active band characteristic of cathepsin V. There was no difference in active cathepsins detected between K, L and S transfected cells compared to vehicle controls (**A**). HEK293T cells were transfected with cathepsin S, cystatin C or both cathepsin S and cystatin C mRNA. After 4 hours, cells were switched to media containing 50µg/mL DQ Elastin. Fluorescence in the conditioned media was measured after an additional 4 or 20 hours after the addition of DQ Elastin. Cells transfected with cathepsin S mRNA did not produce more fluorescence than untransfected controls. Fluorescence was decreased in cystatin C transfected cells, particularly at 20 hours.

5.4 Discussion

Cathepsin and cystatin C specific mRNA successfully induced overexpression of cathepsins S, L, V, K^{V171A}, K^{L253V} and cystatin C in both HEK293T cells and MDA-MB-231 breast cancer cells. Cathepsin V was the only overexpressed enzyme to produce a

detectable zymography signal, indicating the other overexpressed cathepsins were not catalytically active in transfected cells. Overexpressed cystatin C reduced elastinolysis in conditioned media from transfected cells, confirming its inhibitory potential. Overexpression of cathepsins in MDA-MB-231 breast cancer cells did not cause any changes in off-target cathepsins measured by Western blot.

Based on the success of most of the cathepsin mRNA in producing the target protein, we hypothesize that there is an issue with the cathepsin K^{wt} and K^{L253A} mRNA, rather than a problem with the transfection protocol since the same procedure was used for all mRNA. Our collaborators in the Santangelo lab are producing new batches of these mRNA to determine if there was a problem when producing the previous batch.

Detecting active overexpressed cathepsins in only the cathepsin V mRNA transfected cells is surprising, since DNA transfection has previously yielded active cathepsins K, L, S and V in the same HEK cell line. Further experimental optimization is required to determine why cathepsins produced by these engineered mRNA are not catalytically active. Potential causes could include improper glycosylation of the procathepsins, failure to tag the cathepsins in the Golgi with mannose 6-phosphate, or insufficient mannose 6-phosphate receptor proteins to traffic the cathepsins to the endolysosomal pathway.

6 PURIFICATION OF RECOMBINANT CATHEPSINS

6.1 Introduction

Measuring the kinetics of cathepsin cannibalism presents a unique array of challenges compared to measuring more conventional enzyme substrate reactions. First, cathepsins are promiscuous proteases and cannibalistic interactions between two cathepsins can be bidirectional, requiring mutated versions of these enzymes that are resistant to cannibalism or unable to cleave substrates, including other cathepsins. Cathepsins also act as both the substrate and enzyme in these reactions, necessitating a larger amount of enzyme than would normally be required to measure kinetic rates. These challenges motivated the generation of plasmids for active site dead and cannibalism resistant mutant plasmids of cathepsins K, L and S and the generation of stable HEK293T cell lines that either express cathepsins constitutively with a cytomegalovirus promoter, or in response to tetracycline treatment due to repressor proteins inhibiting transcription [103].

The cathepsin plasmids also feature a 6x-Histidine tag (His-Tag), allowing for purification through immobilized metal ion affinity chromatography [112]. The His-Tag reversibly binds to nickel cations allowing for the removal of off-target proteins before eluting the target protein with an Imidazole buffer, which outcompetes for the His-Tag, releasing it from the nickel column. Optimizing a method for producing large amounts of recombinant wildtype and mutant cathepsins would allow for further investigations into the exact mechanisms of cathepsin cannibalism.

6.2 Materials and Methods

6.2.1 Generation of cathepsin overexpressing cell lines

HEK293T cells were stably transfected with plasmids coding for 6x-Histidne tagged wildtype and mutant cathepsins as previously described [103]. HEK293T cells were transfected with linearized plasmids containing cathepsin S or cathepsin L DNA under control of a T-REx (Invitrogen) expression system consisting of a cytomegalovirus promotor controlled by two tetracycline repressor sites. Treatment with tetracycline releases repressor proteins from the repressor sites, allowing transcription to occur.

6.2.2 Cell culture and media collection

HEK293T^{S+} and RAW264.7 cells were cultured in DMEM containing 10% FBS, 1% L-Glutamine and 1% Penicillin Streptomycin. Cathepsin S overexpression was induced in HEK293T^{S+} cells by adding tetracycline to the culture media for 6 hours before changing the cells to serum-free media also containing tetracycline for an additional 24 hours. Conditioned serum-free media was collected from RAW264.7 cells after 24 hours.

6.2.3 Nickel chelating resin protein purification

G Biosciences Nickel Chelating Resin was used to purify 6x-Histidine tagged cathepsins from conditioned media and lysate of cathepsin overexpressing HEK293T cells according to the manufacturer's protocol for native proteins. Resin was washed with water and binding buffer in a spin column before incubation with media or lysate for 60 minutes at 4°C. Resin was collected in a spin column, washed with wash buffer and then target proteins were eluted in multiple fractions with an elution buffer containing 250mM Imidazole.

6.2.4 Western blot and SDS-PAGE

Equal volumes of purified media samples were mixed with reducing SDS-PAGE loading buffer, briefly boiled, loaded and separated on 12.5% polyacrylamide gels. For Western blots, proteins were transferred to nitrocellulose membranes, blocked then probed with primary antibodies for cathepsin S (R&D Systems), 6x-His tag (Invitrogen) or cathepsin L (R&D Systems). For SDS-PAGE, gels were stained with Coomassie Blue and destained.

6.2.5 Multiplex Cathepsin Zymography

Equal volumes of purified media samples were mixed with nonreducing SDS-PAGE loading buffer and cathepsin zymography was performed as previously described [72, 73].

6.2.6 Cathepsin Activation

Cathepsins in conditioned media were activated through incubation in an assay buffer containing 20mM sodium citrate, 10mM dithiothreitol, 25 μg/ml dextran sulfate, pH 4.6 at 37°C [113].

6.2.7 Desalting

Samples were desalted using a Zeba Spin Desalting Column (7kDA MWCO, 10mL, ThermoFisher Scientific) according to manufacturer's recommendations. The desalting column was washed with water, then samples were loaded and recovered via centrifugation. Remaining sample was collected by adding a water stacker and centrifuging again.

6.3 Results

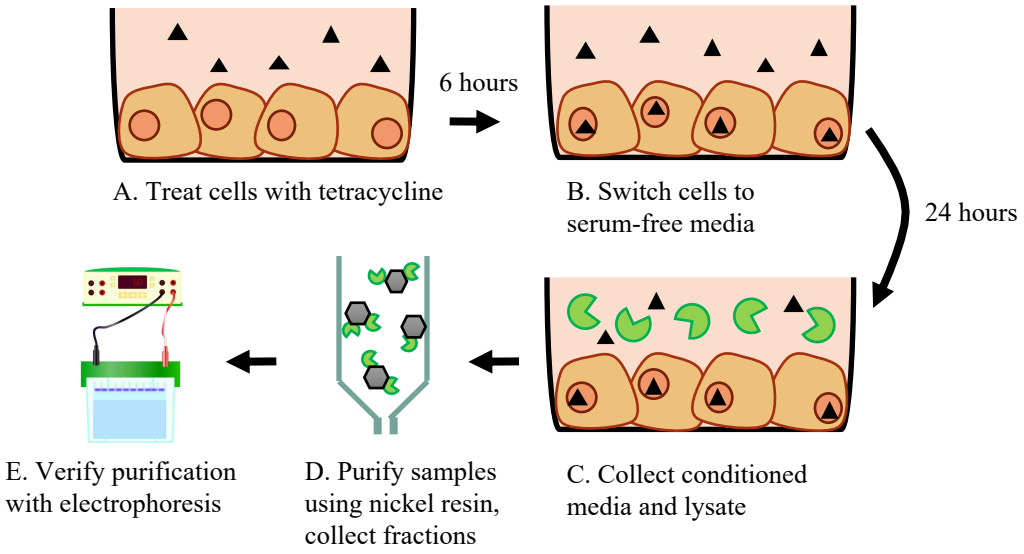

Figure 6-1: Overview of protein purification process from cathepsin overexpressing cells. HEK293T cells were previously transfected with a 6xHistidine human cathepsin S plasmid containing tetracycline repressor sequences. Cells were stimulated with tetracycline for 6 hours in complete medium (A) before being switched to serum-free media containing tetracycline (B). Conditioned serum-free media was collected after 24 hours (C), run through a nickel resin column, and purified proteins were eluted in fractions using imidazole (D). Purification was validated using SDS-PAGE, Western blot and cathepsin zymography (E).

6.3.1 Purification of cathepsin S using nickel chelating resin

Engineered HEK293T cells were grown to confluence in two 15cm cell culture dishes. Cathepsin S expression was induced by adding tetracycline to the culture media for six hours before switching the cells to serum-free media containing tetracycline for an additional 24 hours (Figure 6-1). After 24 hours, media and lysate were collected and purified using 1 mL of nickel chelating resin in a spin column and analyzed using Western blot (Figure 6-2A) and SDS-PAGE (Figure 6-2B). Pro and mature cathepsin S bands were clearly detectable in the unpurified lysate and lysate runoff, indicating very little of the target protein was purified by the column. Faint procathepsin S bands in the lysate fractions are further evidence of the poor purification from cell lysate. Clear pro and mature

cathepsin S bands were clearly detectable in the media purification fractions, indicating a much more effective purification from the conditioned media compared to cell lysate. The vast majority of the purified enzyme was found in the first 3 elution fractions. A faint procathepsin S band was also detectable in the media runoff, indicating that 1 mL of nickel chelating resin was insufficient to capture all of the cathepsins in the cell culture media. SDS-PAGE revealed a 37 kDa band characteristic of procathepsin S in the first purified media fraction. The presence of multiple bands in the media runoff, but only a band at 37 kDa in the media fraction suggests the purification successfully removed the off-target proteins from the sample.

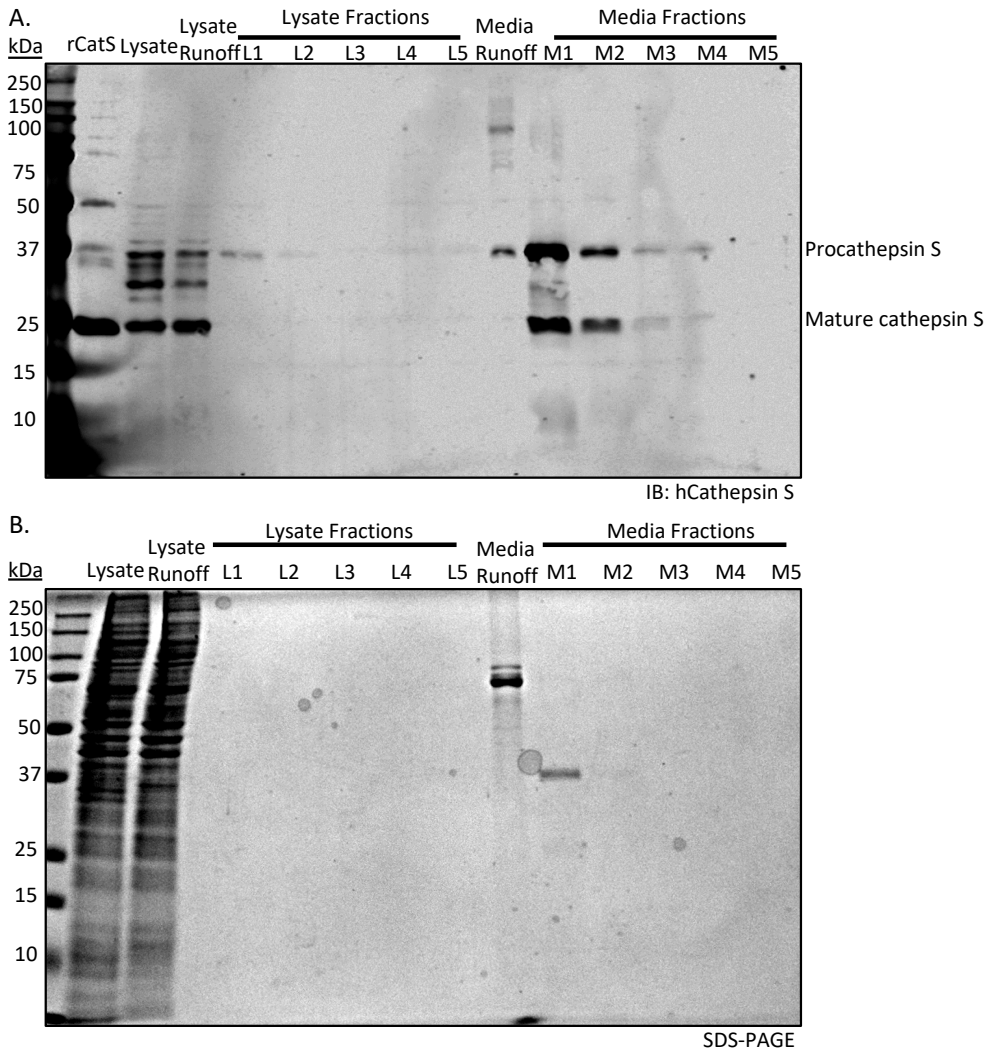

Figure 6-2: Purified pro and mature cathepsin S is detectable in media elution fractions. Cell lysate and conditioned media were collected from tetracycline induced cathepsin S overexpressing HEK293T cells, purified on a nickel resin column, eluted with 250mM imidazole buffer five times and analyzed by Western Blot (A) and SDS-PAGE (B). Pro and mature cathepsin S was detectable in the unpurified lysate and lysate runoff. Procathepsin S was weakly detected in only the first lysate elution fraction. Pro and mature cathepsin S were detectable in the first four elution fractions, and procathepsin S was weakly detectable in the media runoff. SDS-PAGE detected multiple bands in the media runoff, the strongest at 75 kDa and 37 and 25 kDa bands in the first media fraction elution, characteristic of pro and mature cathepsin S, respectively.

Presence of the His-Tag, and absence of cathepsin L, in the purified samples was confirmed by Western blot (Figure 6-3). In the first two purified fractions, strong His-Tag bands were detectable at 37 and 25 kDa, characteristic of procathepsin S and mature

cathepsin S, respectively. 37 kDa bands were also detectable in the lysate, first purified lysate fraction and media runoff, indicating again that lysate purification is ineffective, and a greater volume of nickel resin is required to capture all of the cathepsin S produced in the media. Cathepsin L has previously been detected in HEK293T cells. Due to the similar protein structure and function between cathepsins S and L, we wanted to confirm none of the purified media samples contained cathepsin L. Purified media samples were analyzed by Western blot, probing with an antibody for human cathepsin L. Cathepsin L was detectable in the recombinant control, but not in any lysate or media samples.

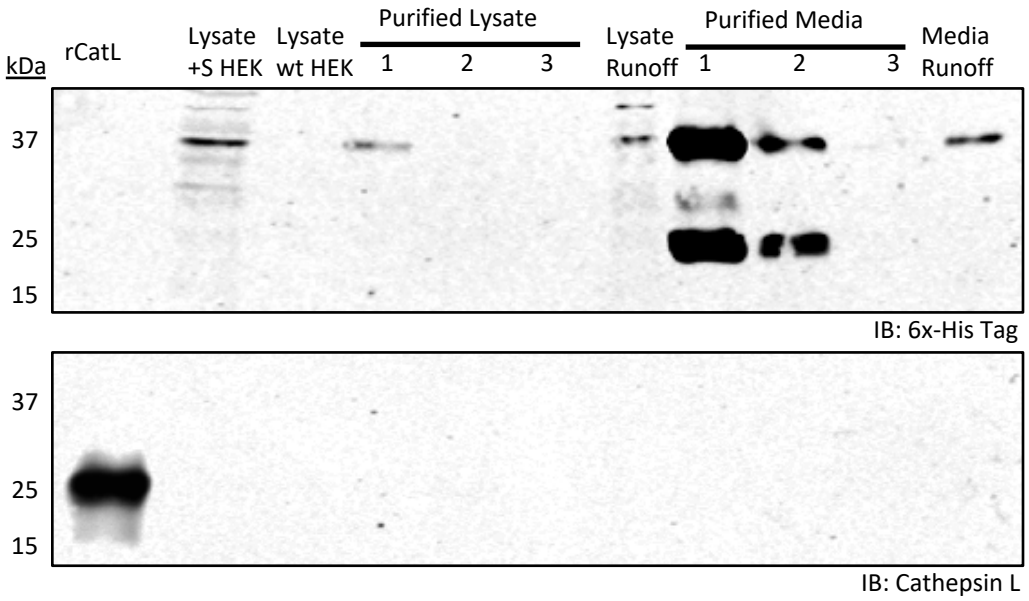

Figure 6-3: Purified cathepsin S is 6x-His tagged and no cathepsin L is present in the purified media. Cell lysate and conditioned media from cathepsin S overexpressing cells were collected and analyzed by Western blot. An antibody for 6x-Histidine revealed strong bands at 37 and 25 kDa, consistent with pro and mature cathepsin S. Cathepsin L was undetectable in the cell lysate and media before and after purification.

To assess whether the purified cathepsins were active and capable of degrading substrate, purified fractions were analyzed with cathepsin zymography (Figure 6-4). Multiple strong bands were detectable in the unpurified lysate, but in no other runoff or

purified fraction. A lack of zymography signal in any of the purified samples, despite Western detectable mature cathepsin S bands, suggests the purification process inactivates mature cathepsins, preventing them from degrading substrate following purification.

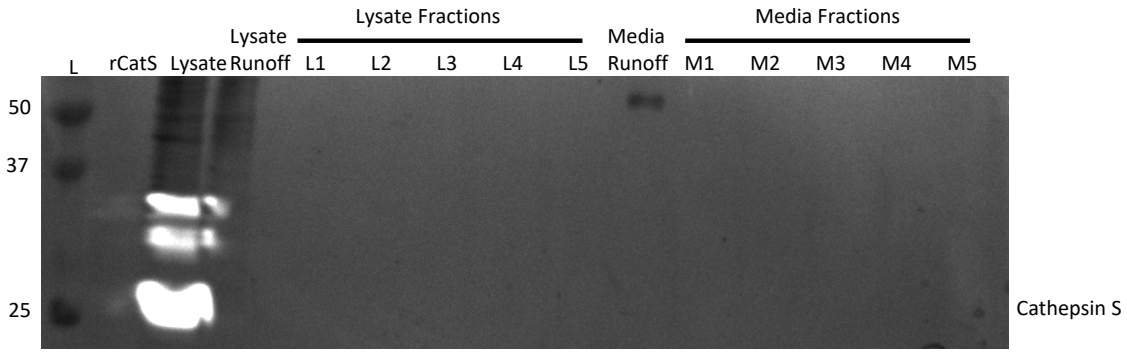

Figure 6-4: Purified cathepsin S is not catalytically active. Cell lysate and conditioned media from cathepsin S overexpressing cells were collected and analyzed by multiplex cathepsin zymography. Strong active bands were detectable in the lysate of cathepsin S overexpressing cells, but no bands were visible in any of the purified fractions or runoff.

Purification from HEK293T cells overexpressing His-Tag cathepsins L and K were also attempted. Cathepsin L was successfully purified from the cell media and analyzed by Western blot (Figure 6-5). Strong procathepsin L bands were detectable in the purification fractions, though the presence of procathepsin L in the media runoff suggests a greater volume of nickel resin should be used when purifying cathepsin L. Cathepsin K was detectable in unpurified cell lysate, but undetectable in media purification fractions (data not shown).

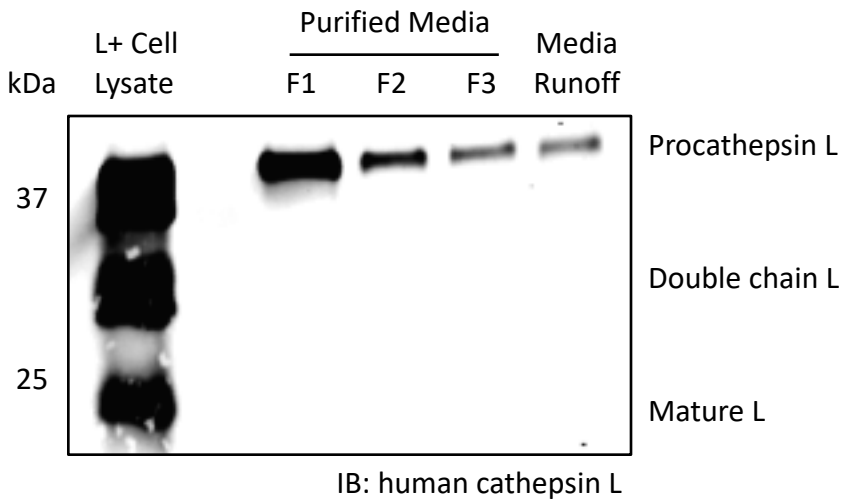

Figure 6-5: Cathepsin L is detectable in purified media samples. Cell lysate and conditioned media were collected from tetracycline induced cathepsin L overexpressing HEK293T cells, purified on a nickel resin column, eluted with 250mM imidazole buffer three times and analyzed by Western Blot. Procathepsin L was clearly detectable in the cell lysate, all three purification fractions and media runoff. Double chain and mature cathepsin L were detectable in cell lysate but not in the media fractions or runoff.

6.3.2 Attempted activation of purified cathepsin S acidic dextran sulfate buffer

Cathepsins are better able to tolerate neutral pH in their unactivated proforms than in their mature forms, following propeptide cleavage [114]. While cathepsin S is better able to retain its structure and activity at neutral pH than cathepsins L or K, the lack of detectable zymography signal indicates that some combination of factors during the purification process rendered the Western detectable mature cathepsin S in the purified media inactive. Cathepsin S activation proceeds much more quickly in an environment mimicking its physiologic endolysosomal niche, namely low pH and reducing environment [29]. Multiple activation protocols, including pepsin mediated activation and incubation at acidic pH in the presence of dextran sulfate, were tested on the purified cathepsin S samples, but none were able to generate zymography detectable active cathepsin S (not shown) [22, 115].

The elution buffer used has a high salt concentration (500 mM NaCl). The high ionic strength of this solution could be inhibiting the autocatalytic and protein-protein binding events necessary for successful activation of cathepsin S. Purified media from cathepsin S HEKs was desalted by flowing through a Zeba spin desalting column and samples were analyzed by Western blot (Figure 6-6). Procathepsin S was faintly detectable in purified samples that were not desalted, and samples desalted then activated for 0 or 30 minutes. No mature cathepsin S was detectable in any of the media samples (Figure 6-6A). Desalted samples were centrifugally concentrated approximately 20-fold then mixed 1:1 with 2x activation buffer and analyzed by Western blot (Figure 6-6B). Pro and mature cathepsin S bands were clearly detectable in the lysate, unactivated and activated samples. However, there was no conversion from pro to mature cathepsin S over 5 minutes in activation buffer. Additional experiments found no zymography detectable cathepsins after 5 or 30 minutes, and no increase in mature cathepsin S after 30 minutes of activation (data not shown).

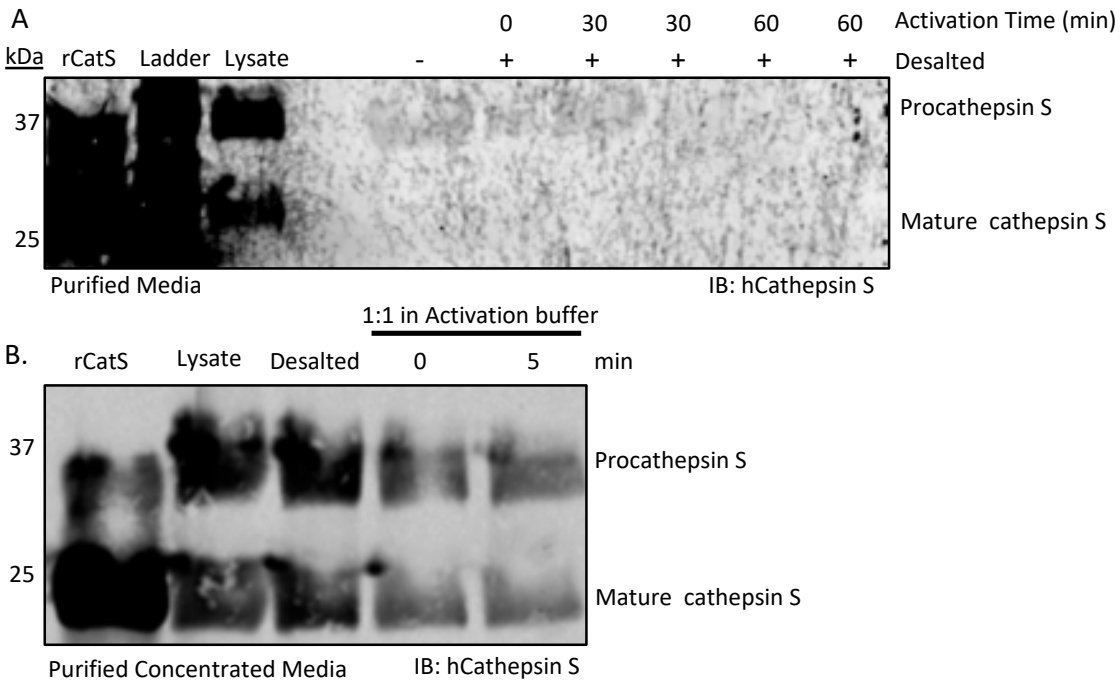

Figure 6-6: Desalting media from cathepsin overexpressing cells does not affect activation. Conditioned media from HEK293T^{S+} cells was purified and desalted prior to incubation with cathepsin activation buffer and analyzed by Western blot (A). Procathepsin S was faintly detectable prior to desalting and undetectable by 60 minutes, with no mature cathepsin S detectable at any time point. Purified media was centrifugally concentrated prior to desalting and activation and analyzed by Western bot (B). Pro and mature cathepsin S were detected in the desalted and activated media. However, no significant transition from pro to mature cathepsin S was detectable in 5 minutes.

To determine which activation protocols can successfully activate cathepsins secreted in media, conditioned media was collected from RAW264.7 macrophages, a murine cell line that produces large amounts of cathepsins, and cathepsin S overexpressing cells. Pepsin mediated activation was not able to generate zymography detectable cathepsins in either RAW or HEK samples. Incubation in pH 4.6 citrate buffer containing 25 µg/ml dextran sulfate greatly increased detectable active cathepsins after 30 minutes of incubation at 37°C (Figure 6-7). In contrast, cathepsin S overexpressing HEK media showed no increase in active cathepsins following incubation in the same buffer, suggesting the HEK cathepsin S does not activate under the same conditions as the RAW

macrophage cathepsins, requiring the optimization of a new activation protocol for cathepsin S secreted by engineered HEK cells.

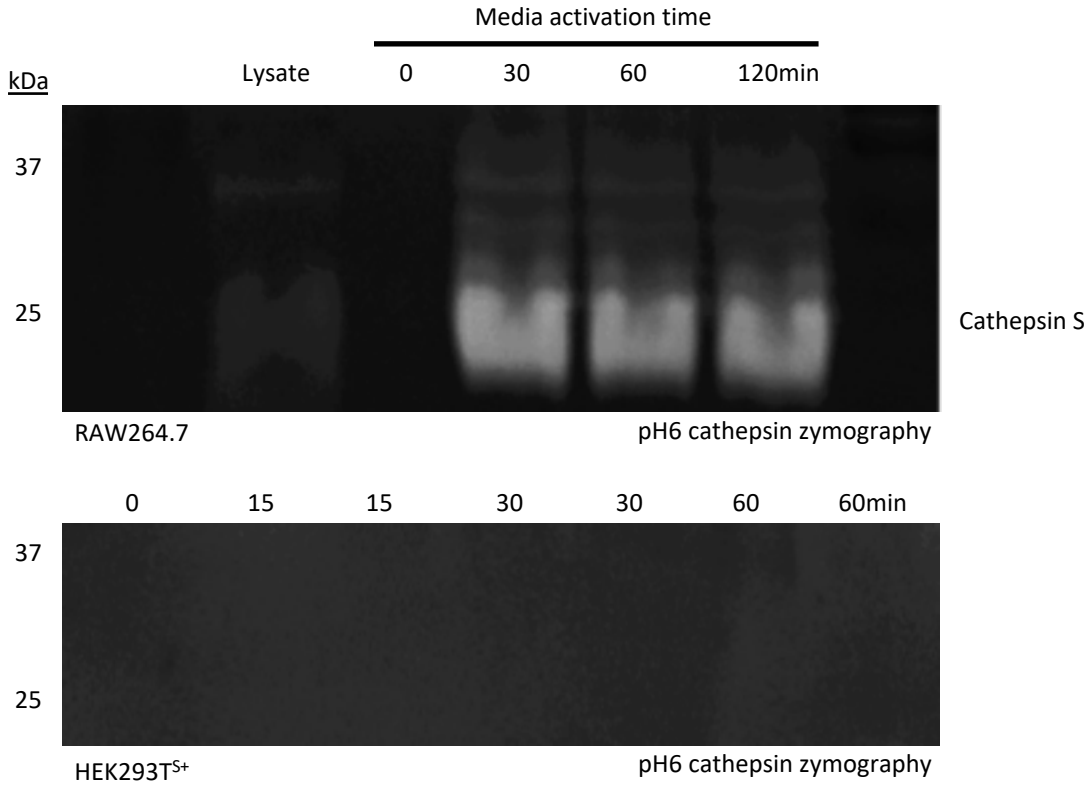

Figure 6-7: Cathepsin activation buffer incubation increases active cathepsin S in media from RAW264.7 cells, but not from cathepsin overexpressing cells. Conditioned media was collected from RAW264.7 (A) and HEK293T^{S+} (B) cells and incubated with a cathepsin activation buffer for up to 120 minutes at 37°C. Prior to activation, no active cathepsins were detectable in RAW264.7 media. After 30 minutes of incubation, a strong active cathepsin S band was detectable, and persisted through 60 and 120 minutes. Conditioned media from the HEK293T^{S+} cells produced no active bands regardless of activation time.

6.4 Discussion

Procathepsins S and L were successfully purified from conditioned media of engineered HEK293T cells using immobilized metal chromatography, moving the lab closer to in-house production of custom cathepsin mutants. No off-target proteins were detectable in the purified fractions, suggesting the protocol effectively concentrates only the target recombinant proteins. However, the resulting purified cathepsin was not

catalytically active, and attempts to activate the procathepsins through incubation in acidic buffers containing other proteases or dextran sulfate were unsuccessful, limiting the experimental usefulness of the purified enzyme. Despite high intracellular concentrations of the target cathepsins, purification from cell lysate was shown to be ineffective, likely due to the large amount of off-target proteins in cell lysate compared to media.

Development of a protocol to produce, purify and activate recombinant wildtype and mutant cathepsins would facilitate kinetic experiments to unravel the exact mechanisms at play during cathepsin cannibalism. Successful production of active site dead and cannibalism resistant mutants is necessary to measure cannibalism parameters accurately because cannibalism is a dynamic process where both enzymes have the potential to act on the other as substrate. Active site mutants reduce this complexity because they would strictly act as substrates for other cathepsins. Cannibalism resistant mutants can similarly be used to measure rates of cannibalism while also providing evidence for the exact location and mechanism of cathepsin-on-cathepsin binding and cleavage.

7 DYNAMIC MODEL OF PROTEASE STATE AND INHIBITOR TRAFFICKING TO PREDICT PROTEASE ACTIVITY IN BREAST CANCER CELLS

7.1 Introduction

Recent advances in breast cancer treatment have drastically increased survival rates for early stage cancer patients, but metastatic stage 4 breast cancer remains extremely difficult to treat with a 5-year survival rate of only 22 percent [116]. Cysteine cathepsins have been implicated in cancer, with recent research revealing integral roles of cysteine cathepsins in promoting tumor growth and metastasis [26, 42]. Cysteine cathepsins are attractive pharmaceutical targets for the prevention of tumor metastasis since they not only are produced by the cancer cells themselves, but also participate in the accessory stromal cells of the tumors and the secondary metastatic sites, as with bone metastases. While some cathepsin inhibitors have shown efficacy in treating bone metastases, to-date, no cathepsin inhibitors have been approved for clinical use due to side effects, particularly skin lesions [43, 44]. The mechanisms responsible for these side effects are unknown, but could include inhibition of off-target proteases, inhibition in off-target tissues, or disruption of proteolytic networks necessary for normal cathepsin regulation.

Cathepsin expression and activity is regulated by a complex proteolytic network consisting of a wide variety of proteases, specific endogenous inhibitors and biological substrates [30]. Cathepsins are translated as inactive procathepsins, which must be cleaved, either autocatalytically or by another protease, before they can become active mature enzymes capable of degrading substrates[18,37]. Changes to the localization, expression or activity of the proteases responsible for cleaving procathepsins can prevent or stimulate

cathepsin activity. Cathepsins and other proteases are also capable of degrading mature enzymes, preventing them from degrading substrates. We have previously reported that cathepsin S is capable of degrading cathepsin K, even in the presence of extracellular matrix substrates, resulting in an overall decrease in collagen proteolysis [70]. Given the contrary ability of proteases to both increase and decrease the activity of other proteases, it is unsurprising that inhibiting a single protease can cause unexpected side effects in clinical trials and in laboratory settings.

While pharmaceutical cathepsin inhibitors display strong, nanomolar enzyme specificity *in vitro*, inhibitor specificity can decrease substantially in intracellular environments. Many early cathepsin inhibitors were lysosomotropic, weakly basic, lipophilic molecules capable of passing through lipid membranes until they are protonated in acidic compartments such as the lysosome [117]. Lysosomotropic inhibitors, such as balicatib, have shown high intracellular potency, but poor specificity, preventing them from passing clinical trials [118, 119]. Emerging non-basic, non-lysosomotropic cathepsin inhibitors have shown improved *in vivo* selectivity and efficacy in clinical trials for osteoporosis and metabolic bone disease, but even these more specific inhibitors have resulted in unexplained side effects in patients [32].

Unravelling the specific interactions of this network is difficult with only experimental methods due to the number of molecules involved and limitations of tracking specific molecules intracellularly. Mathematical modeling of cathepsin kinetics has been instrumental in characterizing the unique responses of recombinant proteins to substrates and inhibitors *in vitro* [64, 65].

This study seeks to uncover the post-translational regulatory mechanisms responsible for the previously observed increase in active cathepsin S following cathepsin inhibitor treatment in breast cancer cells. Additionally, this work expands the focus from the MDA-MB-231 triple-negative breast cancer cell line, to other breast cancer and macrophage cell lines, as well as primary human tumor and patient-matched normal tissue. This work utilizes a combination of experimental and mathematical modeling approaches to explore the mechanism responsible for previously observed increase in active cathepsins following cathepsin inhibitor treatment. The development of accurate mathematical cathepsin-inhibitor models will be instrumental to designing cathepsin inhibitor regimens capable of suppressing tumor growth and metastasis in breast cancer patients.

7.2 Materials and Methods

7.2.1 Cell culture

MDA-MB-231, MCF-7, and MCF-10A cells were incubated in DMEM (Lonza) medium with 10% FBS, 1% l-glutamine, and 1% non-essential amino acids and at 37°C with 5% CO_2. Thp-1 macrophages were derived from Thp-1 monocytes cultured in RPMI-1640 (Lonza) medium with 10% FBS, 1% L-glutamine and 1% Penicillin-Streptomycin. Thp-1 monocytes were differentiated into macrophages by treatment with 0.1uM phorbol 12-myristate 13-acetate (PMA) for 24 hours, followed by 24 hours in fresh medium without PMA before E64 (trans-Epoxysuccinyl-L-leucylamido(4-guanidino)butane) treatment.

7.2.2 Human tissue

Human cancer and normal breast tissue specimens were from National Disease Research Interchange (Philadelphia, PA).

7.2.3 Multiplex Cathepsin Zymography

Total protein concentration of cell lysate and homogenized tissue was measured using the Pierce Micro BCA Protein Assay (Thermo Scientific). Equal quantities of protein were prepared in nonreducing SDS-PAGE loading buffer and cathepsin zymography was performed as previously described [84, 85]. Briefly, samples were loaded into a gelatin-embedded polyacrylamide, lysate proteins were separated by molecular weight during electrophoresis. Following electrophoresis, cathepsins in gels were renatured before being incubated overnight in cathepsin-specific assay buffer at 37°C. The gels were then stained with Coomassie Blue, then destained. White bands indicate the presence of active cathepsins capable of degrading gelatin embedded in the polyacrylamide gel.

7.2.4 Systems modeling of cathepsin dynamics

Previous studies of cathepsins in the MDA-MB-231 cell line have shown that cathepsin S and L are localized in distinct compartments [50]. Accordingly, we have assumed that cathepsin S and L do not directly interact with each other, and the network diagram (Figure 7-4) shows the assumed structure of both systems. The schematic shows relationships between the four cathepsin species, as well as the presence of the inhibitor, E64, which has been reported to irreversibly bind to mature active cathepsin S or L.

Based on the diagram shown in Figure 7-4, we have derived a system of ordinary differential equations (ODE) that describe cathepsin dynamics using simple mass action:

$$\frac{\mathrm{d[Pro]}}{dt} = \mathrm{k_{prod}} - \mathrm{k_{act}} \cdot \mathrm{[Pro]} \quad (1)$$

$$\frac{\mathrm{d[Mat]}}{dt} = \mathrm{k_{act}} \cdot \mathrm{[Pro]} - \mathrm{k_{inact}} \cdot \mathrm{[Mat]} - \mathrm{k_{deg}} \cdot \mathrm{[Mat]} - \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} \quad (2)$$

$$\frac{d[\mathrm{Mat_{inact}}]}{dt} = \mathrm{k_{inact}} \cdot \mathrm{[Mat]} - \mathrm{k_{degI}} \cdot \mathrm{[Mat_{inact}]} \quad (3)$$

$$\frac{\mathrm{d[MatE64]}}{dt} = \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} - \mathrm{k_{degE64}} \cdot \mathrm{[MatE64]} \quad (4)$$

$$\frac{\mathrm{d[E64]}}{dt} = \mathrm{k_{in,E64}} - \mathrm{k_{out,E64}} \cdot \mathrm{[E64]} - \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} \quad (5)$$

This system of equations contains five dynamic variables corresponding to the four cathepsin species and E64, as well as nine kinetic parameters describing the fluxes within the system. Based on our zymography and Western blot data we aimed to identify values of the kinetic parameters that are in agreement with our experimental findings. Parameterization of the dynamic model was performed using a steady state approach, and a detailed description is provided in the Supplementary Methods. During parameter estimation, we investigated two different definitions for the relationship between our experimental measurements and the species of the ODE system. The first definition assumes that all cathepsin species are detectable by Western blot (*WB*), while only active mature cathepsin, *[Mat]*, is detectable through zymography (*Zymo*), as defined by Eqs. 6 and 7.

$$WB = \mathrm{[Mat]} + \mathrm{[Mat_{inact}]} + \mathrm{[MatE64]} \quad (6)$$

$$Zymo = \mathrm{[Mat]} \quad (7)$$

The second definition still assumes all cathepsin species are detectable by Western blot, but differs from definition 1 in that active mature cathepsin and E64-bound cathepsin are both assumed to be detectable through zymography (Eqs. 8 and 9)

$$WB = [\mathrm{Mat}] + [\mathrm{Mat_{inact}}] + [\mathrm{MatE64}] \quad (8)$$

$$Zymo = [\mathrm{Mat}] + [\mathrm{MatE64}] \quad (9)$$

Dynamic simulations were performed based on the steady state kinetic parameters and the system of ODEs defined by Eqs. 1 – 5. The rate of influx and outflux for E64 were approximated to ensure that the assumed steady state E64 concentrations were reached within 24 hours. Additionally, the kinetic parameters were scaled to ensure the system reached a steady state within 24 hrs. Numerical simulations were performed using the deSolve package in the R statistical language [120].

7.2.5 Statistical analysis

Zymography band intensities were quantified by densitometry in ImageJ and data were normalized by dividing each band intensity on a gel by the intensity measured from cells or tissue treated with 0 μM E64 (setting this group to 1 across all gels). Two-tailed Student's t-test with two-sample equal variance was performed on all statistical analysis.

7.3 Results

7.3.1 Incubation with E64 reduces amount of active cathepsin L in MDA-MB-231 cells, but increases the amount of active cathepsin S

Whether the response of increased active CatS and decreased CatL after incubation with the broad cathepsin inhibitor E64 was specific to MDA-MB-231 cells or could occur in other cell types was yet to be established. To test this, the invasive estrogen receptor positive MCF-7 [121], non-transformed, non-invasive breast epithelial MCF-10A [122], and invasive, triple negative MDA-MB-231 [123] cell lines, were incubated with increasing concentrations of E64 for 24 hours then lysed and prepared for cathepsin zymography to determine amounts of active cathepsins. E-64 incubation increased active cathepsin S but led to a decreased amount of active cathepsin L in MDA-MB-231 cells (Figure 7-1A). E64 appeared to slightly increase active cathepsin S in MCF-7 and MCF10A cells, but this increase was not statistically significant when averaged across three biological replicates. Neither MCF-7 nor MCF10A cells expressed active cathepsin L detectable by zymography, regardless of E64 treatment (Figure 7-1B and Figure 7-1C).

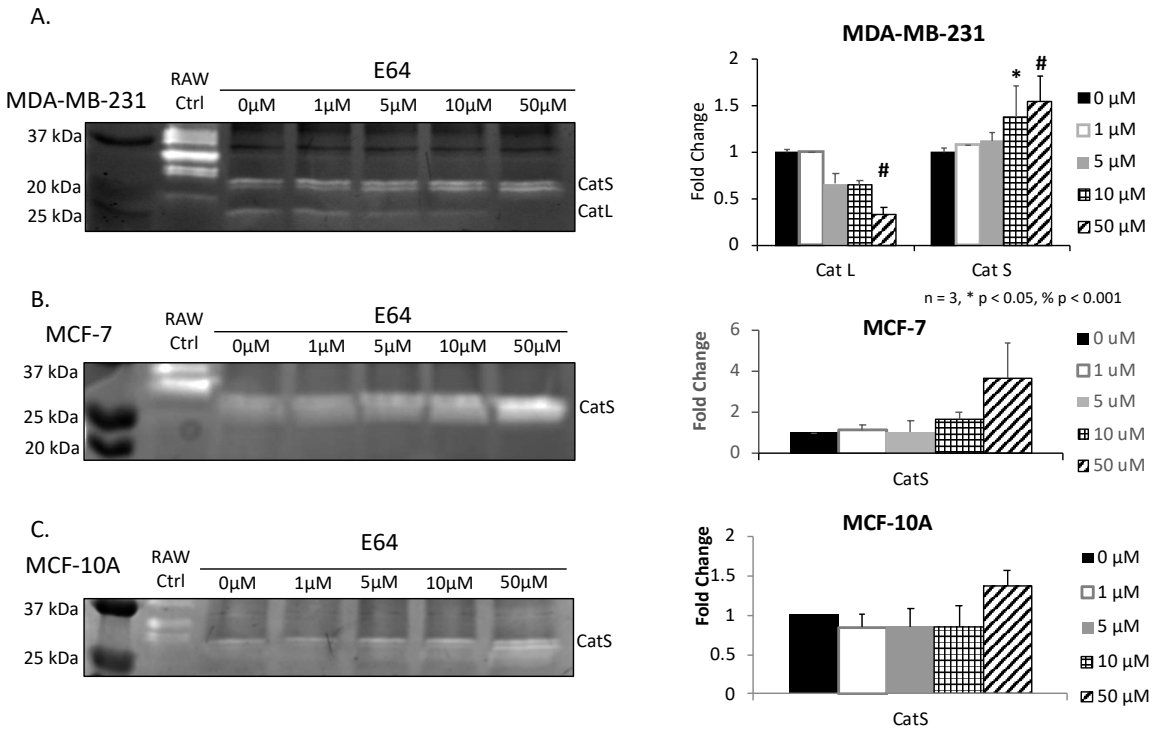

Figure 7-1: Incubation with E64 reduces amount of active cathepsin L in MDA-MB-231 cells, but increases the amount of active cathepsin S. MDA-MB-231 (A), MCF-7 (B), and MCF-10A (C) cells were grown to confluence prior to incubation in the presence of 0, 1, 5, 10 and 50 μM E64 for 24 hours. Cells were lysed and equal amounts of protein were loaded for multiplex cathepsin zymography. Densitometry was used to quantify the intensity of the active cathepsin S and L signals and shown in the graphs on the right for each cell type. There was an almost 50% increase in active cathepsin S, but a 75% loss of active cathepsin L in the MDA-MB-231 cells (n=3, *p<0.05, %p< 0.001). Neither MCF-7 nor MCF-10A had detectable levels of active cathepsin L.

7.3.2 E64 treatment of human normal and cancerous breast tissue yields variable responses of active cathepsins

After observing increased active cathepsins following E64 treatment in one breast cancer cell line, the next step was to determine if this same response would occur in primary human tissue. Breast tissue contains multiple cell types, but we hypothesized that human breast tumors would exhibit similar increases in active cathepsin S following E64 treatment. Patient-matched tumor and normal breast tissue were acquired, cut into small

pieces, and incubated in increasing concentrations of E64 for 24 hours. Normal and cancer tissues were homogenized in lysis buffer prior to being assayed by cathepsin zymography (Figure 7-2A); zymography results were quantified by densitometry (Figure 7-2B).

In the absence of E64, active cathepsins S and L were detectable in cancer and normal tissue (Figure 7-2A), consistent with previous publications indicating cathepsin involvement in breast cancer [124]. The amount of active cathepsin was highly variable between patient samples, with four of the samples showing active cathepsin V signal in addition to previously detected active cathepsins L and S. For the most part, E64 increased cathepsin S in normal tissue and decreased cathepsin L. However, in cancerous tissue, E64 caused a significant decrease in cathepsins L, S, and V, with some amount of patient to patient variability.

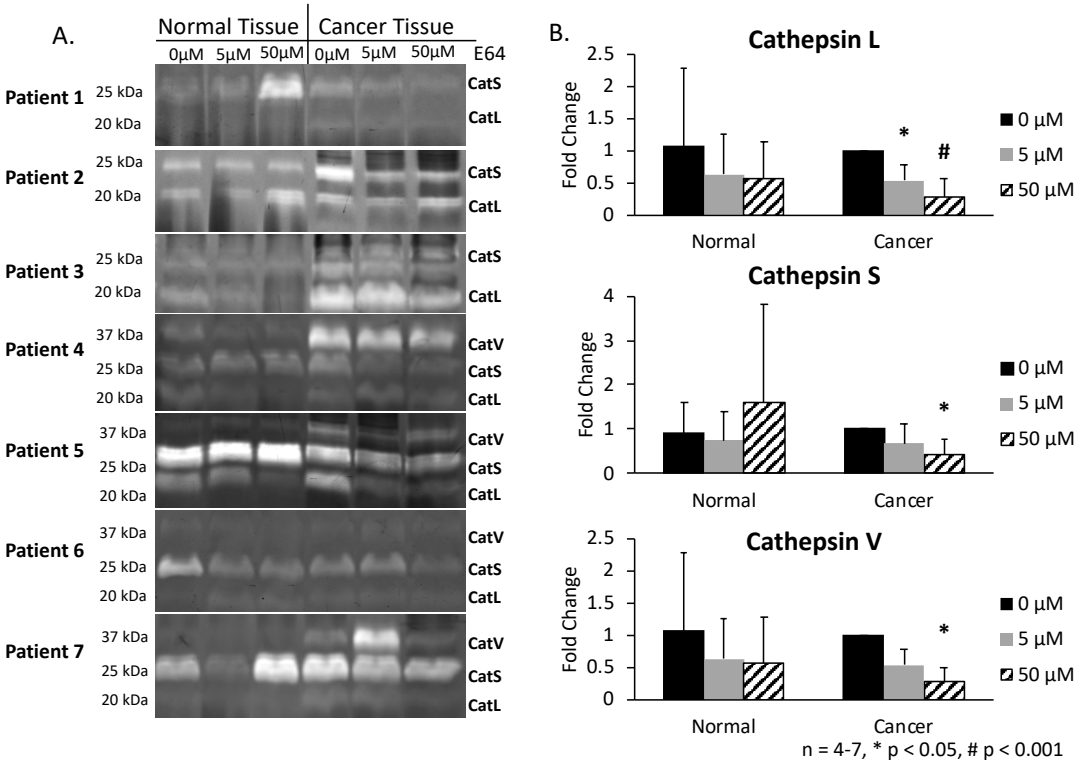

Figure 7-2 Treatment with E64 stimulates higher amounts of active cathepsin S in some normal, but not tumor tissue. Excised, patient-matched normal and tumor breast tissue was collected from seven human patients, and equally sized thin pieces were incubated in media

containing 0, 5, or 50 μM of E-64 for 24 hours, then lysed, and equal amounts of proteins were loaded for multiplex cathepsin zymography (A). Active cathepsins L and S were detectable in all patient samples, and active cathepsin V was detectable in four samples. Active cathepsins and response to inhibitor treatment was highly variable between patients, but active cathepsins L, S and V were significantly decreased by over 50% in the cancer, but not the normal tissues (n = 4-7, * $p < 0.05$, # $p < 0.001$) (B).

7.3.3 Mass action model of cathepsin inhibition, unable to explain experimental results when assuming E64 inhibits zymography signal

A mathematical model of cathepsin L and S interactions with E64 in MDA-MB-231 cells was constructed to explain the counterintuitive responses to E64. Reactions in the model include: synthesis/production of procathepsins, activation from procathepsin to the mature, active cathepsin, degradation, inactivation, and inhibition by E64 (Figure 7-4). Based on this diagram, we designed a system of ordinary differential equations (ODE) using simple mass action functions (Eqs. 1 – 5) that describe cathepsin dynamics. This system of equations contains five dynamic variables corresponding to the four cathepsin species and E64, as well as nine kinetic parameters describing the fluxes within the system. Based on our zymography and Western blot data we aimed to identify values of the kinetic parameters that are in agreement with our experimental findings.

The model was fit to Western blot and cathepsin zymography data collected from MDA-MB-231 cells treated with 0, 10, or 50 μM E64 for 24 hours under the assumption that zymography detected mature, but not inactive, cathepsins (Mat), since enzymes must be active to generate a signal. The procathepsins, mature cathepsin, inactivated and inhibited cathepsin forms would all be detected by Western blots. Assumptions of the model also were that cathepsin L was in the cytoplasm and cathepsin S was intravesicular as shown previously [50]. Parameterization of the dynamic model was performed using a steady state approach. We assumed that the E64 concentration in the lysosomal

compartment, which is where cathepsin S is localized, reaches a steady state concentration equal to the specified media concentration after 24 hours. This assumption was made because time course measurements using 50 μM E64 in MDA-MB-231 cells found the amounts of zymography detectable cathepsins L and S had largely plateaued by 24 hours . Steady state concentrations in the cytosol, which is where cathepsin L is localized, were scaled based on an approximation that the cytosol volume is 35 times larger than the lysosomal volume. Based on these assumptions, and the proposed system of ODEs (Eqs. 1 – 5), steady state equations were defined for the three tested concentrations of E64: 0 μM, 10 μM, and 50 μM. This system of equations includes seven unknown steady state kinetic parameters, as well as steady state concentrations of cathepsin species that are not directly measurable based on available experimental data. Detailed equations (Equ 6 – 28) and parameter estimation steps are included in Appendix 1.

The steady state equations were reduced to a system of six equations (Eqs. 29 – 34) with six cathepsin-related kinetic parameters as unknowns. Eqs. 29 – 32 define the relationship between the kinetic parameters and the steady state concentrations, while Eqs. 33 and 34 serve as constraints connecting the steady state concentrations to the experimental data.

$$[\mathrm{Mat}]_{10\mu M} - [\mathrm{Pro}]_{0\mu M} \cdot \frac{k_{act}}{k_{inact} + k_{deg} + k_{E64} \cdot [\mathrm{E64}]_{10\mu M}} = 0 \tag{29}$$

$$[\mathrm{Mat}]_{50\mu M} - [\mathrm{Pro}]_{0\mu M} \cdot \frac{k_{act}}{k_{inact} + k_{deg} + k_{E64} \cdot [\mathrm{E64}]_{50\mu M}} = 0 \tag{30}$$

$$\mathrm{k_{act}} \cdot [\mathrm{Pro}]_{0\mu M} - [\mathrm{Mat}]_{0\mu M} \cdot \left(\mathrm{k_{inact}} + \mathrm{k_{deg}}\right) = 0 \tag{31}$$

$$k_{inact} \cdot [Mat]_{0\mu M} - k_{degI} \cdot [Mat_{inact}]_{0\mu M} = 0 \quad (32)$$

$$[Mat]_{10\mu M} + [Mat_{inact}]_{10\mu M} + [MatE64]_{10\mu M} - WB_{10\mu M} = 0 \quad (33)$$

$$[Mat]_{50\mu M} + [Mat_{inact}]_{50\mu M} + [MatE64]_{50\mu M} - WB_{50\mu M} = 0 \quad (34)$$

Parameter estimation was initially performed for cathepsin S and L, where the squared residuals for Eqs. 29 – 34 were minimized with respect to the six kinetic parameters. These equations define a nonlinear root finding problem composed of six equations and six unknowns, where an exact solution cannot be determined analytically. Therefore, a numerical approximate solution was obtained using a newly developed optimization algorithm for black-box problems by identifying kinetic parameter values that minimized the sum of the squared residuals for Eqs. 29 – 34 [125]. The applied algorithm is surrogate-based, and utilizes a special class of polynomials to approximate the relationship between the optimization variables and the objective function. For this problem the optimization variables were the parameter values and the sum of the squared-residuals of the steady-state equations was the objective function that was being minimized. The algorithm starts with an initial set of samples that fall on a Smolyak sparse-grid, and then iteratively samples the parameter space by using the polynomial approximation to predict were the best samples lie and then updates the approximation following the collection of new samples. Local optimization of the polynomial approximation is also used in identifying which samples to collect next (i.e. the parameter combinations to evaluate), and criteria based on the improvement of the polynomial approximation are used for determining convergence. Based on this model, parameter estimation for cathepsin L

proved successful with the identification of a solution with low residuals, but for the case of cathepsin S no reasonable solution was found (Table 7-1).

Table 7-1: Residuals for initial model, assuming E64 bound cathepsin S is undetectable by zymography

	Residuals Squared (μM^2)	
Equation	Cathepsin L	Cathepsin S
29	3.5E-09	7.836
30	1.6E-06	8.302
31	6.2E-08	1.014
32	1.5E-08	3.0E-04
33	1.3E-09	8.978
34	2.2E-06	7.150
Sum of Squared Residuals	3.9E-06	33.279

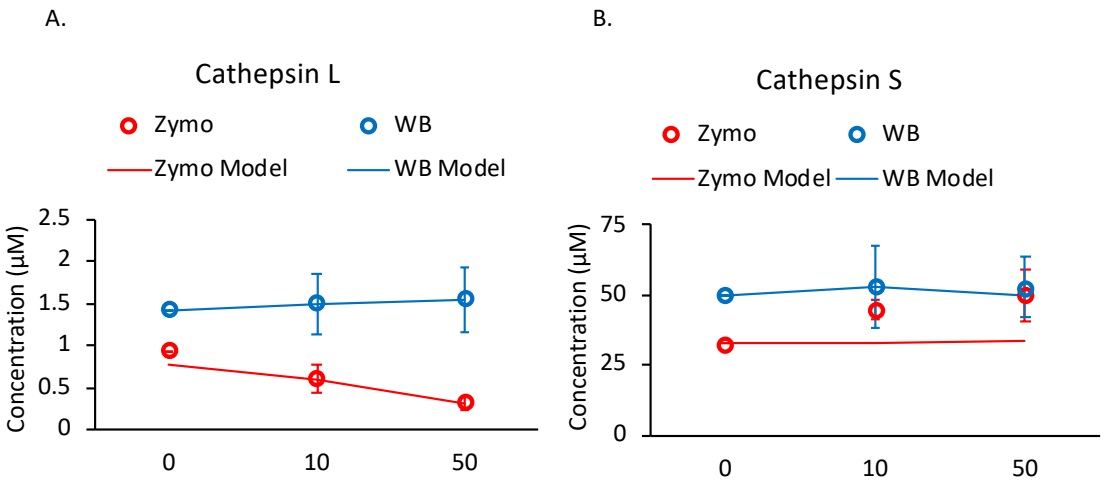

Figure 7-3: Initial model accurately fits cathepsin L but not cathepsin S zymography data. Cathepsin L **(A)** and cathepsin S **(B)** model fits to experimental Western blot (WB) and zymography (Zymo) data at different dosages of E64. The cathepsin L model fits the zymography and Western blot data. The cathepsin S model is able to fit the Western blot data, but does not predict the increase in zymography signal following E64 treatment seen in MDA-MB-231 cells.

The model was able to simulate the behavior of E64 reducing active cathepsin L, but not the unexpected result of E64 treatment increasing active cathepsin S (Figure 7-3A&B). A number of alternative reaction scenarios were considered including E64 affecting rates of activation and degradation of cathepsins S and L. These repreliminarily modeled to determine if they could fit the data and better describe the system in action, but none of these models were able to recapitulate the results of increased active cathepsin S as seen in the zymography results. Transcriptional regulation of cathepsin S was not considered because E64 treatment did not affect cathepsin S mRNA or protein in MDA-MB-231 cells [50]. After exhausting these alternative explanations, it was necessary to reassess our initial assumption that cathepsins are only detectable by zymography when they are not bound by the cathepsin inhibitor E64.

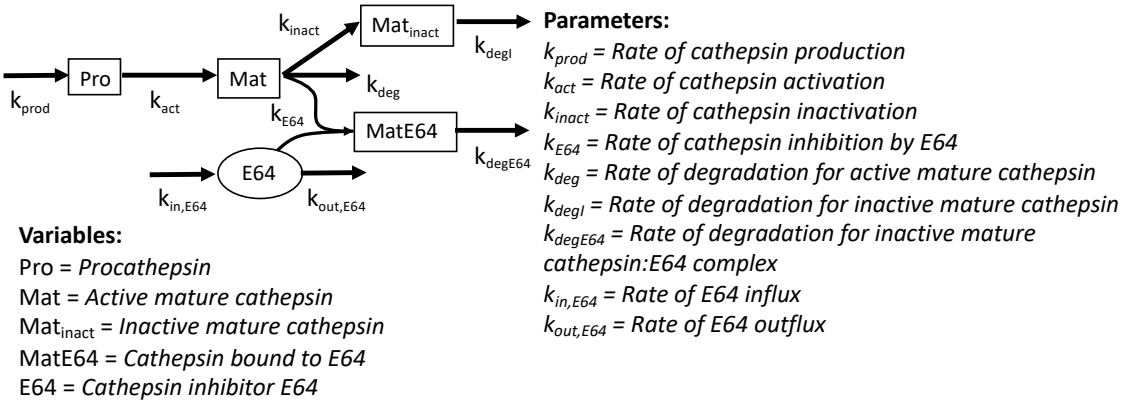

Figure 7-4: Cathepsins L and S are located in different intracellular compartments, leading to different responses to inhibitors with different entry mechanisms. Cathepsins L and S are translated in the rough endoplasmic reticulum, trafficked through the Golgi apparatus to endolysosomal vesicles bound for lysosomes (A). Cathepsin S remains in endolysosomal vesicles, while cathepsin L enters the cytoplasm (B). E64 cannot freely pass through the plasma membrane, and enters the endolysosomal vesicles, where it can bind cathepsin S, before leaking into the cytoplasm where cathepsin L is available to bind (C). E64d is a derivative of E64 that contains an ethyl ester which allows it to cross the cell membrane. Once in the cytosol, the ester is hydrolyzed by cytoplasmic esterases, trapping it in the cytoplasm where it can bind to cathepsin L (D). Lysosomotropic inhibitors, modeled on lipid hormones capable of freely moving across the plasma and lysosomal membranes where they can bind cathepsins L and S respectively (E). Schematic representation of cathepsin E64 model architecture. Cathepsins are translated in pro form, activated to mature form, then inactivated, degraded, or bound by E64.

7.3.4 E64 does not block cathepsin S zymography signal under cell lysate conditions

E64 is considered an irreversible cathepsin inhibitor, due to the formation of a covalent bond between the C2 atom of the E-64 epoxy ring and the active site cysteine of the cathepsin, as well as hydrogen bonding between E-64 and polar residues near the cathepsin active site [93]. However, under the partial denaturing conditions of zymography followed by refolding after the electrophoresis, the cathepsin-E64 complex could potentially be disassociated or shifted into a non-inhibitory conformation. To test the hypothesis that cathepsin S and E64 were uncoupling during the denaturation/renaturation of zymography, MDA-MB-231 cells were cultured as before, with 0 or 50 μM E64 for 24 hours, then collected in zymography lysis buffer. An additional 10 μM of fresh E64 was

added to the cell lysate, assuming additional E64 would bind to free cathepsin L and S in the lysate. MDA-MB-231 cells treated with 50 μM E64 showed decreased active cathepsin L and increased active cathepsin S, consistent with previous results (Figure 7-5A). There was no change in the zymography signal for those samples that had an additional 10 μM E64 added to the lysate in lysis buffer compared to the control samples for neither cathepsin L nor S (Figure 7-5A).

To ensure that 10 μM E64 was sufficient to bind and inhibit cathepsins, a zymogram was run in parallel, but was treated differently after electrophoresis and cathepsins were renatured; the gel was incubated in cathepsin assay buffer containing 10 μM E64. Incubation in assay buffer with 10 μM E64 completely abrogated the cathepsins' zymography signal, even in the RAW264.7 macrophage lysate used as a positive control (Figure 7-5B). Densitometry of the active cathepsin zymography signal is shown (Figure 7-5C). The ability of E64 to inhibit cathepsins when added to the assay buffer, but not when added to lysate, suggests the zymography process disrupts the inhibitory properties of E64, possibly through conformational changes to the cathepsin active site brought on by sodium dodecyl sulfate (SDS), detergents in the lysis buffer, basic pH of lysis buffer or electrophoresis buffers. This potential unbinding of E64 under non-reducing SDS-PAGE conditions was not expected because E64 has been classified as an irreversible covalently binding inhibitor of the active site cysteine [93], but this led to reconsideration of the experimental data to which the computational model was being fit.

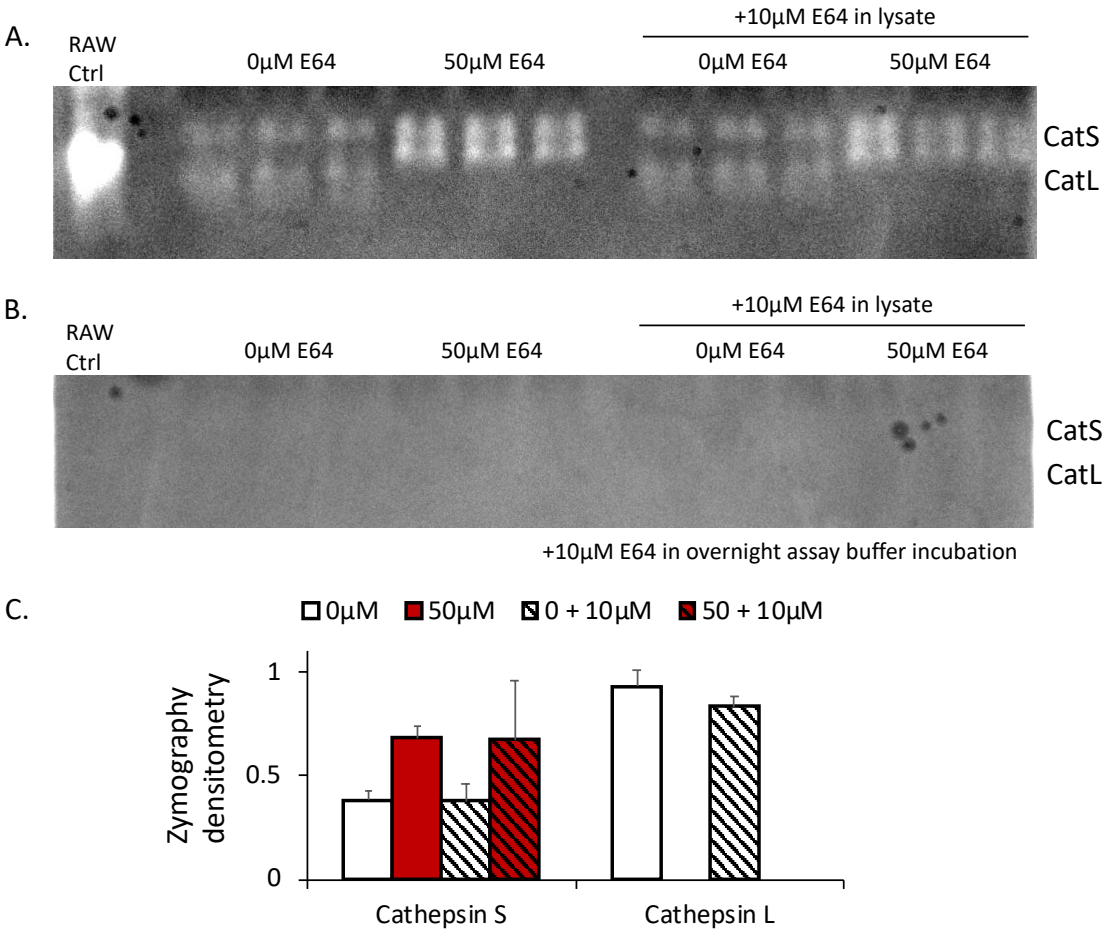

Figure 7-5: E64 does not block cathepsin S zymography signal under cell lysate conditions. MDA-MB-231 cells were cultured with 0 or 50 μM E64 for 24 hours then the cells were lysed and lysate collected from each of the incubation conditions. Aliquots of each lysate were then incubated with 0 or 10 μM E64 for 10 minutes to allow binding of E64 to active cathepsins. Then equal protein amounts from these samples were prepared for cathepsin zymography as usual (A), or 10 μM fresh E64 was added to the overnight incubation step during cathepsin degradation of the embedded gelatin (B). There was no observed difference in active cathepsin zymography signal between lysates treated with additional, fresh E64 or not, but incubation with E64 during the degradation completely blocked all active cathepsin zymograms signal. Densitometry was used to quantify the zymography signal (n=3) (C).

7.3.5 Correcting for cathepsin S-E64 zymography signal significantly improves model fit, parsing inhibited cathepsin from uninhibited cathepsins, and active from inactive conformations

The mathematical model was revised to incorporate the ability of cathepsin S bound by E64 in living cells to generate a zymography signal, as shown schematically in Figure 7-6A. Model fitting with these new assumptions generated parameters with much lower residuals for the cathepsin S parameters (Table 2), recapitulating the experimental results garnered from Western blots and zymography (Figure 7-6C&D). Steady state simulations of the model predicted dose dependent accumulation of cathepsin-E64 complexes for both cathepsins L and S (checkerboard segments of graphs), but only cathepsin S-E64 complexes, not cathepsin L-E64 complexes, were allowed to contribute to zymography signal (Figure 7-6B). Taken together, the red signals indicate zymography signal generating species with a direct relationship with E64 concentration for cathepsin S but an inverse relationship for cathepsin L.

Table 7-2: Residuals for revised model, assuming E64 bound cathepsin S is detectable by zymography

	Residuals Squared (μM^2)
Equation	Cathepsin S
6	6.9E-14
7	1.1E-13
8	2.3E-13
9	1.0E-13
10	6.3E-14
11	1.8E-13
12	2.0E-12
13	2.8E-13
14	2.0E-14
15	4.8E-13
16	4.8E-13
19	1.5E-11
35	5.2E-13
21	1.8E-11
36	2.9E-13
Sum of Squared Residuals	3.8E-11

Model parameters governing cathepsin production (k_{prod}), activation (k_{act}), inactivation (k_{inact}), and degradation (k_{deg}) were estimated to be several orders of magnitude greater for cathepsin L than cathepsin S (Table 7-3), indicating cathepsin L production and turnover to be greater than that of cathepsin S in MDA-MB-231 cells.

Table 7-3: Parameters for final MDA-MB-231E64 model

Parameter	Cathepsin L	Cathepsin S
kProd	4.67E-02	1.40E-05
kAct	9.34E-04	2.75E-07
kInact	3.91E-04	5.71E-08
kDeg	1.31E-03	3.76E-07
kDegI	4.77E-04	1.34E-07
kE64	1.81E-03	7.05E-09
kDegE64	9.24E-04	1.81E-07

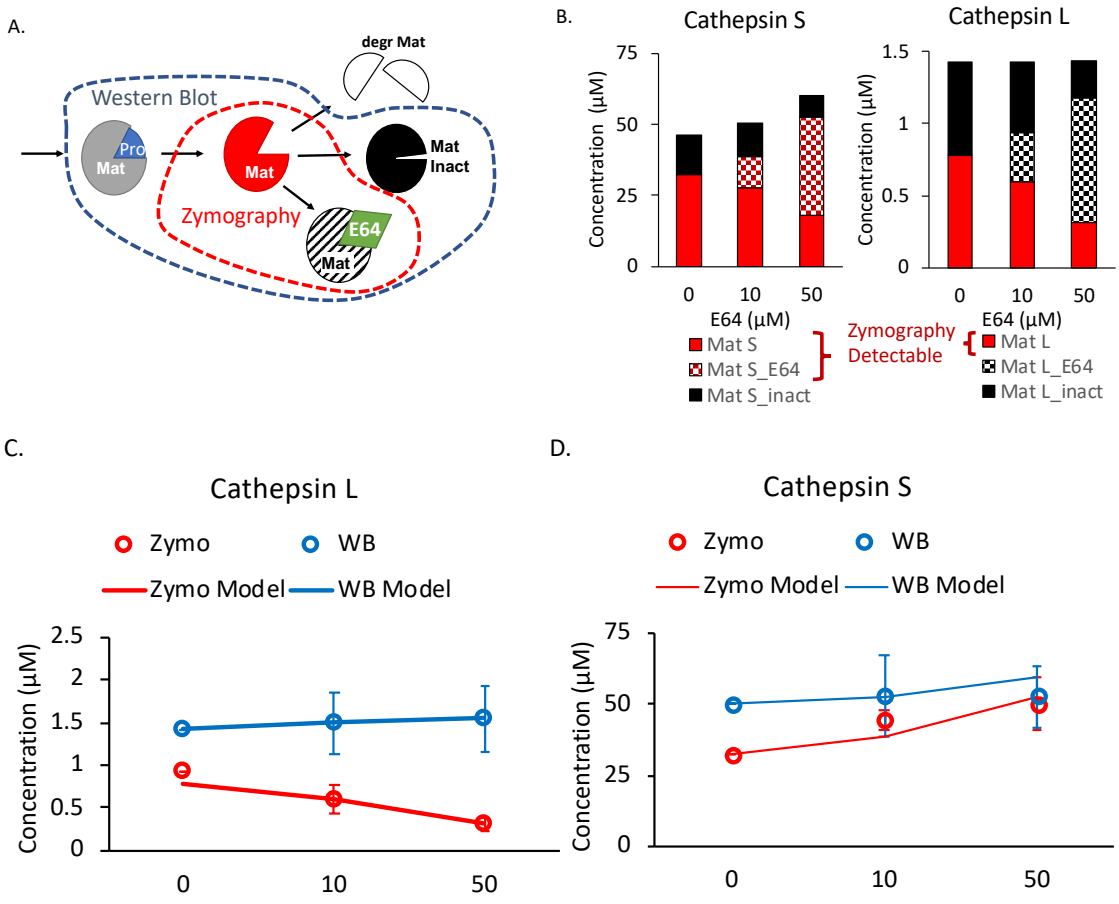

Figure 7-6: Correcting for cathepsin S-E64 zymography signal significantly improves model fit, and parses bound from unbound states and active from inactive conformations.

Revised model architecture assumes cathepsin S is detectable by Western blot when in its pro state, free mature state, inactive state or when bound by E64 and detectable by zymography when in its free mature state or when bound by E64 (A). Model predictions for enzyme species at steady state for cathepsins S and L. The model predicts that E64 treatment increases cathepsin bound E64 complexes and decreases free cathepsins in a dose dependent manner. E64 is also predicted to decrease non-E64 inactivated cathepsin L by over 50% while non-E64 inactivated cathepsin S is less affected (B). The model is able to recapitulate experimental results for cathepsin L **(C)** and cathepsin S **(D)**.

7.3.6 *A priori* predictions of dynamic responses of cathepsins L and S to different classes of cathepsin inhibitors

Following the successful construction of the steady state cathepsin L and S model, we were interested in constructing a dynamic model with utility to describe the kinetics of different classes of cathepsin inhibitors binding to and inhibiting cathepsins over time. The model includes extracellular, lysosomal (cathepsin S) and cytoplasmic (cathepsin L) compartments, and the class of cathepsin inhibitor determined their ability to move between these compartments: membrane impermeable cathepsin inhibitor such as E64, membrane permeable inhibitor such as E64d, and a lysosomotropic inhibitor, which made up a number of early cathepsin inhibitors that advanced to clinical trials but were unsuccessful [119, 126, 127] (Figure 7-7A). The steady state cathepsin E64 models were used to parameterize time course models to predict cathepsin dynamics following inhibitor treatment. Inhibitor treatment was modeled as a bolus of 10 μM at time point 24 hours, and models were simulated until 48 hours.

Membrane impermeable E64 was modeled as an endocytosed inhibitor taken up by the cell into endolysosomal compartments that fuse with lysosomes, then able to bind to cathepsin S. The model assumes the small molecule E64 is able to leak out of the endolysosomal compartment into the cytosol in a similar manner to other lysosomal

contents that have been shown to move into the cytoplasm in cancer cells [128]. Once these minute amounts of E64 are in the cytoplasm, they are able to bind to cathepsin L present in the cytoplasm, as demonstrated by immunocytochemistry previously [50]. The dynamic model predicted that cells treated with E64 would accumulate cathepsin L-E64 and cathepsin S-E64 complexes, peaking 24 hours after treatment, resulting in elevated cathepsin S zymography signals and diminished cathepsin L zymography (Figure 7-7B), as was observed experimentally (Figure 7-1). Cathepsin L was predicted to reach peak inhibition in under 1 hour, compared to cathepsin S, which did not reach its maximum zymography signal until 12 hours after inhibitor treatment.

For E64d, the membrane permeable inhibitor due to addition of an ethyl ester that is removed by cytosolic esterases, effectively trapping E64d in the cytoplasm, the results were different. The model predicted E64d would have little effect on cathepsin S, since it was located in the endolysosomal compartment. However, E64d would be extremely effective at inhibiting cathepsin L due to its ease of crossing the membrane and entering the cytoplasm. According to the simulation, E64d was more effective than E64 at inhibiting cathepsin L (22% inhibition for E64 and nearly 100% inhibition for E64d), but the inhibition subsided more quickly since the greater influx of inhibitor resulted in a more rapid depletion of free inhibitor (Figure 7-7C), compared to the limited release of E64 into the cytoplasm that resulted in a much slower return to baseline for free cathepsin L. As predicted, cathepsin S responded less to E64d than E64 (22% increase in zymography signal for E64 compared to 7% increase for E64d). The predicted greater inhibitory capacity of E64d compared to E64 on cathepsin L and lack of response in cathepsin S are consistent with experiments using E64d in MDA-MB-231 cancer cells, which found 1μM

E64d inhibited intracellular cathepsin L at a 1μM dose, while having no effect on cathepsin S at any dosage [50].

The third class of lysosomotropic inhibitors were modeled as moving freely across lipid membranes and through compartments, from extracellular to cytosolic to endolysosomal. Lysosomotropic inhibitors showed rapid binding to cathepsins S and L to form complexes. The magnitude of inhibition of cathepsin L by the lysosomotropic inhibitor (85%) was greater than for E64, the membrane impermeable inhibitor (22%). The lysosomotropic inhibitor provoked the largest increase in cathepsin S zymography signal (43% for lysosomotropic vs. 22% for E64). These simulations are consistent with studies of lysosomotropic inhibitors that have shown high efficacy in binding target and off-target cathepsins and side effects caused by these potent inhibitors [117, 118, 126].

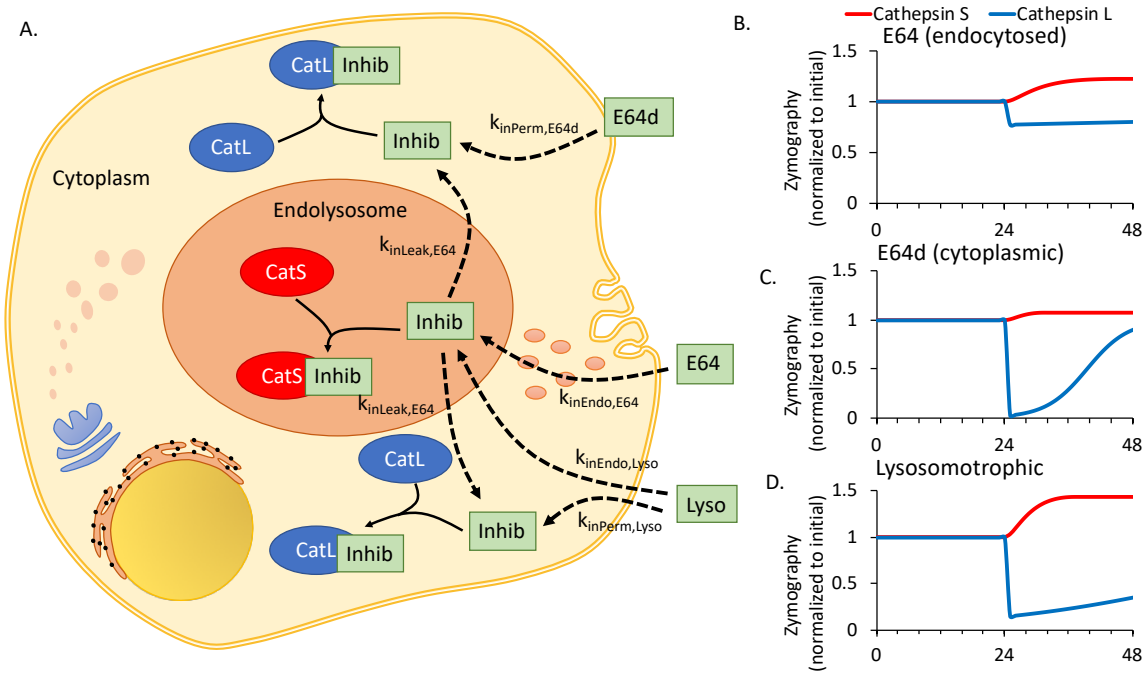

Figure 7-7: Dynamic model predicts different classes of cathepsin inhibitors provoke divergent responses by cathepsins L and S. Schematic representation of dynamic models of different inhibitors. Cathepsin S and L are confined to the endolysosomal and

cytoplasmic compartments, respectively. Inhibitors can enter the endolysosomal compartment through endocytosis, or enter the cytoplasmic compartment through passive diffusion. Inhibitor can also pass from endolysosomal to cytoplasmic through lysosomal leakage (A). Dynamic model predictions of cathepsin L and S states, and predicted zymography and Western blot results following treatment with E64 (B) E64d (C) or a lysosomotropic inhibitor (D). Steady state model parameters were scaled to create a dynamic model of cathepsin inhibition over time. A treatment of 50 μM inhibitor was added at 24 hours, and cathepsin dynamics were predicted for several days following treatment to allow the system to move toward its steady state. E64 was able to bind to both cathepsins L and S through endocytosis and subsequent leakage out of the endolysosomal compartment and into the cytoplasm. Due to its confinement in the cytoplasm, E64d was much more effective in binding cathepsin L than cathepsin S, which is consistent with previous experimental data showing low dosages of E64d can inhibit cathepsin L while not affecting cathepsin S. E64d also inhibited cathepsin L much more rapidly than E64, since it could cross the cell membrane and enter the cytoplasm directly. Lysosomotropic inhibitor bound both cathepsins L and S more rapidly than E64, resulting in loss of free cathepsins L and S.

7.4 Discussion

This work shows treatment with a cathepsin inhibitor can increase different active cathepsins in multiple cell types and primary tissue, suggesting there is a common mechanism among these cell types that should be considered when designing inhibitor treatment regimens. *Ex vivo* tumor biopsy specimens from different patients displayed significant person-to-person heterogeneity in endogenous active cathepsins and in response to inhibitor treatment. The presence of stromal cells, such as tumor associated macrophages, could explain some of the variability observed in the primary tumor samples as macrophages are known to have donor-specific variation in active cathepsin expression [83]. Successful clinical treatment of cancer metastasis with cathepsin inhibitors will require better understanding of the proteolytic interactions, in order to effectively suppress target proteases contributing to metastasis, while avoiding impacting proteases that would provoke unexpected side effects. Additionally, these results underscore the importance of assaying active cathepsins, in addition to total cathepsins, during inhibitor clinical trials.

The models we developed were able to differentiate between inhibitor-bound, active, and inactive cathepsin species, which are difficult to measure experimentally and can confound *in vivo* and *in vitro* experiments. Finally, the dynamic models of different classes of cathepsin inhibitors that have been deployed in clinical trials, demonstrate how inhibitor trafficking and access to subcellular compartments can have drastically divergent effects on active cathepsins located in different intracellular compartments.

Cathepsins are attractive targets for multiple diseases including cancer, osteoporosis and atherosclerosis, but off-target effects and unexpected responses to cathepsin inhibitor treatments have prevented their clinical adoption. This work sought to explain a previously documented non-intuitive response to cathepsin inhibitor treatment in breast cancer cells, a simultaneous increase in active cathepsin S and decrease in cathepsin L following E64 treatment. Several different cells related to breast cancer tumors were experimentally tested to determine the distribution of the non-intuitive E64 treatment response, including different transformed and non-transformed breast cancer cell lines, primary tumor and patient matched normal tissue samples, and a human macrophage cell line.

This work focused on the small molecule inhibitor E64 as a model cysteine cathepsin inhibitor with cross-reactivity. To test the effects of E64 on primary tissue, patient matched normal and tumor samples were incubated with E64. The patient samples displayed diverse endogenous cathepsin signatures, including cathepsins L, S and V. This high degree of proteolytic variability agrees with our previous findings of a high degree of person-to-person variability in endogenous active cathepsin and cystatin expression among monocyte derived macrophages isolated from healthy donors [83]. On average, E64

treatment decreased active cathepsins in tumor samples. However, several normal tissue samples had increased active cathepsins following E64 treatment. These results highlight the difficulty in designing effective cathepsin inhibitor regimens due to high patient variability in endogenous proteases. Additional experimentation will be required to determine the roles of different cells in complex *in vivo* systems, including tumor tissue, which is composed of many different cells including cancer-associated fibroblasts, tumor associated macrophages, other leukocytes, endothelial cells and adipose cells, many of which are known producers of cysteine cathepsins [25]. Additionally, breast cancer cells are known to be highly heterogeneous even within the same primary tumor [129], further underscoring the need for *in silico* tools to study variation in proteolytic inhibitor responses.

To explore the effects of E64 on other cancer associated cells, macrophages were treated with E64 and active cathepsins were quantified. Tumor associated macrophages are abundant in breast tumors, comprising up to 50% of the tumor mass [10]. Cysteine cathepsins produced by tumor associated macrophages have been shown to promote tumor growth and metastasis, angiogenesis, and even to suppress chemotherapy [7, 130]. E64 treatment significantly increased active cathepsin S and active cathepsin V in the Thp-1 human macrophage cell line. Increased active cathepsins in Thp-1 macrophages suggests this effect is conserved across cells of multiple lineages, including epithelial cells derived from the endoderm and macrophages derived from the mesoderm. Additionally, the increase in active cathepsin V indicates this response is not limited to cathepsin S, but could also affect other proteases, particularly other cysteine cathepsins due to the high degree of protein identity shared between family members.

The failure of our original model assumption, that cathepsins bound by E64 are undetectable by cathepsin zymography, to explain the observed increase in active cathepsin S following E64 treatment led us to question this assumption and test it experimentally. Treated lysate with additional E64 did not affect the resulting zymography signal, suggesting E64 loses its inhibitory capacity during non-reducing electrophoresis. We suspect E64 is playing a stabilizing role on cathepsin S during E64 incubation, where free cathepsin S becomes bound to E64 instead of naturally inactivating due to changes in intracellular compartment pH or oxidative conditions. Once E64 binding is interrupted during electrophoresis, the cathepsin S bound by E64 may regain its ability to degrade substrate to generate the elevated zymography signal observed. E64 does not provide the same stabilizing effect intracellularly as it provided to cathepsin S, resulting in loss of active cathepsin L signal, following cellular treatment with E64. This could be caused by differences in the intracellular environments of the two enzymes in MDA-MB-231 cells, since cathepsin L is dominantly in the cytoplasm in these cells while cathepsin S was confined to the endolysosomal compartment. Cathepsin S is also known to tolerate neutral pH more effectively than cathepsin L, which could be quickly denatured in the more neutral pH of cytoplasm or basic pH of lysis buffer, regardless of E64 binding [131].

Dynamic simulations of different classes of cathepsin inhibitors showed the importance of intracellular trafficking mechanisms on the specificity and strength of inhibition. Lysosomotropic inhibitors elicited large responses from both cathepsins L and S due to high cellular uptake and accumulation in the lysosomes, consistent with previous results showing high intracellular potency, but poor selectivity among lysosomotropic cathepsin inhibitors [118, 119, 127].

Previous studies of cathepsin inhibition in rodent models have shown the ability of cathepsin inhibitors to affect the expression and activity of cysteine cathepsins. Cathepsin B, H and L had increased half lives in the livers of rats treated with the potent cysteine cathepsin inhibitor EP-475 (E-64c) [132]. While cathepsin L or cathepsin S protein were not increased following E64 treatment, it is possible that the protein could accumulate over a period of time longer than 24 hours. Treatment with the weaker binding, reversible inhibitor leupeptin elicited increased cathepsin B and other lysosomal hydrolase activity in rat fibroblasts and mouse bone tissue, measured with reporter substrates [133]. However, enzyme activity was abolished when the rat fibroblasts were treated with E64.

This work suggests inhibitors can protect specific enzymes resulting in long-term increases in activity following the unbinding of an inhibitor. Small molecule inhibitor interactions or interference with cystatins, the endogenous protein inhibitors of cathepsins, must also be included. Previous work in our lab has shown cystatin C can elicit the same preservation of active cathepsin S, while interestingly failing to inhibit cathepsin L [50]. Cathepsin stability is greatly affected by binding substrates and effector molecules such as glycosaminoglycans, the presence of which can increase cathepsin K half-life at neutral pH from 7 to 190 minutes [65]. Similarly, binding to E64 could preserve cathepsin S activity, creating a reservoir of active cathepsins capable of degrading substrate following inhibitor unbinding. Our model results suggest even broad-spectrum inhibitors can have diametrically opposed enzyme-specific impacts, preserving or inhibiting activity depending on properties of the enzyme and its intracellular location. The exact effector molecules and mechanisms responsible for this differential response to cathepsin inhibitor binding will be the subjects of future studies.

8 CONCLUSIONS AND FUTURE DIRECTIONS

8.1 Major Findings

In this dissertation, we took important steps forward in developing experimental and computational tools to predict proteolytic dynamics in multiple cell types involved in breast cancer metastasis. We also identified and explored unexpected responses to perturbations of those networks, including cathepsin inhibitor treatment and cathepsin cannibalism.

We were able to determine the limits of detection of cathepsin zymography for measuring active cathepsins, particularly cathepsin V, in human monocytes. Validating this tool will provide valuable insight when applying cathepsin zymography to complex multicellular tissue samples, such as breast tumors, in the future, potentially predicting the macrophage content of primary tumors based on active cathepsin V content. Additionally, we characterized the intracellular and extracellular accumulation rates of cathepsins L and S in human macrophages. These data reveal significant cathepsin-specific differences in protein turnover and accumulation and will inform the development of future computational models of proteolytic dynamics in macrophages. We found that cellular cathepsin L is almost completed depleted following 24 hours of translation inhibition, suggesting a short intracellular half-life due to high cathepsin L turnover or secretion. Interestingly, while cathepsin L was secreted rapidly, it reached a steady state concentration by 8 hours. Cathepsin L reaching a steady state suggests either the enzyme is being degraded extracellularly or that macrophages adjust rates of cathepsin L production or secretion over time, possibly in response to concentrations of intra or extracellular cathepsin L. In contrast, intracellular cathepsin S levels remained relatively

constant following translation inhibition, and extracellular levels increased linearly over the 24-hour time course, though more slowly than the initial secretion of cathepsin L. This points to a lack of extracellular degradation of cathepsin S, or a difference in control over translation and secretion regulation between the two enzymes, since extracellular cathepsin S levels did not reach a steady state after 24 hours.

We also documented an unexpected increase in active cathepsins V and S, and decrease in active L following cathepsin inhibitor treatment in macrophages. We went on to explore a similar phenomenon in breast cancer cells, but differences in the response suggest there may be significant cell type specific effects relevant to cathepsin inhibitor treatment. The macrophages had increased active cathepsin V, which was not detectable in the breast cancer cells, showing this phenomenon is not unique to cathepsin S. The macrophages also had increased pro and mature cathepsin S and pro cathepsin L, while no changes in pro or mature cathepsins L or S were detectable in the breast cancer cells. One hypothesized reason for this difference in response is an increased uptake of cathepsin inhibitor in highly phagocytic macrophages as compared to cancer cells, leading to a large response in pro and mature protein levels resulting in detection in macrophages, but not in cancer cells. Additional experimental and computational investigation will be required to uncover other factors potentially causing this upregulation in response to inhibitor treatment and reasons for its differential behavior based on cell type.

Cathepsin cannibalism had previously been documented *in vitro*, and we observed it in an engineered cellular system mimicking the dynamics between macrophages and in breast cancer cells. We developed and parameterized a mathematical model of cathepsin dynamics to parse the relative contributions of cathepsins K and S to overall cathepsin K

degradation. We identified four distinct families of solutions, which predicted different relative contributions of cathepsins S and K to cathepsin K degradation. We identified cathepsin K activation rate as a critical parameter separating the families into groups where cathepsins were not involved in degradation vs. groups where cathepsins played a significant role in cathepsin K degradation. Repeating the parameter estimation after establishing new bounds on cathepsin K activation rate yielded three new solution families: CatK Dominant, Equal Partners and CatS Featured. While these solutions appeared to fit the data similarly, extrapolation revealed the CatS Featured solutions had the most consistent and least oscillatory behavior, providing evidence for their validity as biological solutions. All of the refined solutions agree that cathepsin K plays at least as large a part in its own degradation as cathepsin S, a surprising result, given the *in vitro* data of the original cathepsin cannibalism studies which found cathepsin S to be more important to cathepsin K degradation. However, cathepsin K having a large impact on its own degradation is consistent with the most recent *in vitro* modeling, which suggests cathepsin K autodigestion could be several times faster than cathepsin S on cathepsin K degradation, at pH 6.

Cathepsin mRNA for protein overexpression was developed and validated in HEK293T and breast cancer cells. Most of the mRNA was able to rapidly induce a large amount of protein overexpression. However, only cathepsin V was found to be catalytically active by zymography. Overexpressed cystatin C was found to reduce DQ elastin degradation in conditioned media of transfected cells, suggesting it has inhibitory potential and can be used in experimental models of cathepsin inhibition. We also tested the hypothesis that overexpression of different cathepsins could induce a change in the

expression of other cathepsin family members. While we saw no change in off-target cathepsins following transfection, it is possible that this result would be different if the mRNA were overexpressing catalytically active enzymes.

Pro and mature cathepsins L and S were successfully purified from the conditioned media of HEK293T cells stably transfected to overexpress cathepsins L or S. Unfortunately, these enzymes were not catalytically active following purification. Experiments with RAW macrophages optimized an activation protocol able to activate macrophage media cathepsins. However, this protocol was unable to activate purified procathepsins or even unpurified cathepsins from the media of overexpressing HEK cells. Additional experiments will be required to optimize an activation protocol for cathepsins in the conditioned media of HEK cells.

Previous work had shown E64 treatment could increase active cathepsins in cancer cells. We extended this work by treating multiple cancer cell lines and primary patient tissue with E64 and measuring active cathepsins. Active cathepsin S was slightly increased in other cancer cell lines and non-cancerous tissue, but significantly decreased in primary tumors. Additionally, primary patient tissue displayed considerable heterogeneity, both in endogenous active cathepsins and response to inhibitor treatment. We developed a computational model of intracellular cathepsin production, inhibition, inactivation and degradation that motivated an experiment to test the ability of zymography to detect E64 bound cathepsins. Experimental results confirmed E64 bound cathepsins can be detected by zymography. This surprising result changed our understanding of the inhibitor response we had previously seen, since the increase in active cathepsin S could now be explained by E64 preserving the activity of cathepsin S similar to how the weak binding cathepsin

inhibitor leupeptin is used in zymography lysis buffer to preserve cathepsin activity [85]. Cathepsin L does not seem to be preserved by E64 binding intracellularly, possibly due to its location in the cytoplasm of MDA-MB-231 cells. We also parameterized a dynamic model capable of simulating cathepsin inhibitors with different intracellular trafficking mechanisms, demonstrating how different cathepsin inhibitor responses can be based on the location of the target cathepsin and trafficking method of the inhibitor.

8.2 Future Research Directions

8.2.1 Proteolytic profiles of human macrophages

The biggest question I have about the human macrophage data is what is causing extracellular cathepsin L levels to grow quite quickly for the first 4-8 hours after CHX removal, then plateau to a steady state for the rest of the time course. Extracellular degradation of cathepsin L by some protease seems likely. Cathepsin S is one possible candidate due to its stability at neutral pH. However, our *in vitro* results have not shown it to be a particularly aggressive degrader of cathepsin L. Another possible candidate is the aspartyl protease cathepsin D, which has been show to degrade inactive cathepsin L *in vitro* [134]. Cathepsin D is more active in the acidic environment of the lysosome, but a low pH pericellular environment could allow it to cleave secreted cathepsin L [109]. Characterizing the secretion rates of macrophages isolated from different primary donors could also be valuable to the development of personalized models to predict responses to cathepsin inhibition among different individuals.

8.2.2 Cathepsin dynamics in multicellular systems

While documenting cathepsin cannibalism in an engineered cellular system was a significant step from *in vitro* studies, it falls short of showing cathepsin cannibalism in a more physiologic system. The obvious next question is where does cathepsin cannibalism occur naturally?

A likely physiologic site of cathepsin cannibalism is the lysosome. The low pH and reducing environment of the lysosomes promote activity of many cathepsins, and a variety of cathepsins and other lysosomal hydrolases are found here. Interestingly, lysosomes exhibit a significant amount of variation in morphology and luminal contents, even among the lysosomes in a single cell [135]. This could complicate efforts to identify effects of lysosomal cathepsin cannibalism and other proteolytic network dynamics in lysosomes, since different versions of these interactions could be playing out differently across the lysosomes of a single cell. Another potential site of cathepsin cannibalism is the cytoplasm during apoptosis.

I believe cathepsin and cystatin C mRNA will prove to be invaluable tools for studying cathepsin proteolytic networks *in vivo* once the details preventing forming active protein are ironed out. Since the transfection efficiency of the mRNA is so high, even among primary cells, we could induce overexpression of cathepsin S in human osteoclasts, potentially degrading cathepsin K, slowing collagen degradation and bone resorption. The creation of a host of cannibalism resistant cathepsin mutant mRNAs could provide experimental evidence in cellular or even animal models of cathepsin cannibalism as a regulatory mechanism.

8.2.3 Molecular approaches to probing the proteolytic network

Cathepsin and cystatin C mRNA is an exciting new tool to perturb proteolytic networks and discover unknown regulatory functions of cathepsins in multiple cell types. Unfortunately, while most of the mRNA produces considerable amounts of protein, the only protein that seems to be catalytically active is cathepsin V. Determining why these mRNA overexpressed proteins are inactive could be a significant undertaking, but the ability to transiently overexpress wildtype and mutant cathepsins with high efficiency in a wide variety of cell types and even in animal models should be well worth the effort. I believe there are two main possibilities for why these overexpressed proteins are inactive, improper folding or incorrect intracellular trafficking. A relatively simple experiment to check if the protein is properly folded is to attempt to activate the overexpressed proteins through the use of a low pH buffer and dextran sulfate, similar to the one used in Chapter 6. However, a more direct method for measuring whether these proteins are folded is circular dichroism and comparing to recombinant protein samples, which are known to be catalytically active. If either of these results confirm the protein is folded, the culprit is likely improper trafficking from the Golgi to endosomes and lysosomes. Potential areas to focus on in this pathway are glycosylation, mannose 6-phosphate tagging and mannose 6-phosphate receptor mediated sorting.

If these issues can be resolved and the mRNA can induce the production of active cathepsins in multiple cell types, a huge variety of experimental possibilities open up. While we were able to attempt to test our hypotheses related to the overexpression of one species of cathepsin affecting the expression or activity of another, the lack of catalytic activity in our overexpressed cathepsins could have prevented the overexpressed enzymes

from impacting the cells, disguising any regulatory effects. Repeating these experiments with catalytically active overexpressed cathepsins could reveal hypothesized regulatory links between different cathepsin family members in cells. Primary cells, especially those that do not divide normally like macrophages, can be difficult to transfect with DNA, but mRNA transfection is much more successful in these cell types [111]. Cathepsin mRNA could be used to better model tumorigenic macrophages by inducing overexpression of different cathepsins in a tumor model. Cathepsin overexpression could also be used to directly compare the relative impacts of different cathepsins on a variety of experimental variables such as tumor invasion, response to cathepsin inhibitors, vascular remodeling, degradation of fibrin clots, and any other experimental model investigating the role of cysteine cathepsins.

Protein overexpression with mRNA has also been shown to work in animal models. In cathepsin knockout mice, another related protease often shoulders some of the work normally done by the knocked out cathepsin. Transfecting knockout mice to overexpress the previously knocked out cathepsin is a potential application to study the relative contributions of different proteases to similar functions *in vivo*. Additionally, the cystatin C mRNA we have seems to successfully inhibit cathepsins, so it could potentially be used to perturb cathepsin systems in cells or animal models. Cystatin C levels can vary considerably between individuals, so cathepsin mRNA would allow us to better recreate those person-to-person differences *in vivo*.

8.2.4 Cathepsin purification and activation

By successfully purifying cathepsins S and L from cell media, I was able to help take the lab one step closer to producing our own mutant and wildtype recombinant

cathepsins on demand. Unfortunately, I was unable to activate any purified samples. After working on this project for some time, I have come to the conclusion that we need protein purification expertise. I recommend actively seeking a post-doctoral researcher with experience setting up a recombinant protein, ideally protease, production and purification system.

The current setup in the lab, with adherent cells grown on cell culture plates or flasks, is time consuming, wasteful of reagents and extremely low throughput. Attempts at purification and activation consumed massive numbers of cells, due to poor tolerance for serum-free media in HEK293Ts. If we ever hope to produce large amounts of cathepsins in-house, we need to invest in higher throughput culturing equipment, such as shake flasks and convert to suspended cells, either re-deriving the cathepsin mutant lines or using transient DNA transfection, which can be very effective when done at scale [136].

I believe a significant factor in our inability to activate purified samples was working with insufficient amounts of protein. Due to our cell culture setup, it took multiple 150mm plates of confluent cells and multiple days to even attempt a single activation protocol. Also based on the effectiveness of activation in RAW macrophage media, I suspect a greater abundance of cathepsins in our HEK media would make it much easier to activate due to the exponential nature of cathepsin activation, where activated cathepsins help cleave the propeptides of other procathepsins.

8.2.5 Cathepsin inhibitors in breast cancer and macrophages

Discovering that E64 bound cathepsins can be detectable by cathepsin zymography was a significant step toward understanding the complicated nature of cathepsin inhibitors in intracellular environments, but there are still many unanswered questions holding back

the use of cathepsin inhibitors to prevent cancer metastasis and treat other pathologies without incurring unexpected side effects. Experiments with cell lysate and recombinant enzymes suggest cathepsins K, L and S retain some zymography detectable activity after E64 treatment despite E64's characterization as an irreversibly binding inhibitor. This is consistent with MDA-MB-231cathepsin S retaining zymography detectable signal, but cathepsin L from the same cells loses its activity following E64 treatment. The conservation of this effect observed in Thp-1 macrophages suggests cathepsin L's loss of activity is tied to fundamental properties of the enzyme rather than a quirk of MDA-MB-231cells. The exact differences between cathepsins L and S that give rise to this variable response to E64 are still open questions.

A potential player in the unexpected active cathepsins seen in E64 treated cells is the lysosomal protease cathepsin B. Cathepsin B has been shown to prevent transcriptional factor EB (TFEB) activation through cleavage of the lysosomal channel protein transient receptor potential cation channel, mucolipin subfamily (TRPML1) in macrophages [92]. Loss of cathepsin B results in activation of TFEB, a transcription factor which promotes lysosomal biogenesis and autophagy through the production of lysosomal hydrolases and membrane proteins. E64 is a broad-spectrum cysteine protease inhibitor capable of inhibiting cathepsin B similarly to cathepsins L and S. Unpublished results from our lab have indicated E64 treatment affects lysosomal distribution in MDA-MB-231cells, likely through a similar mechanism to the cathepsin B mediated TRPML1 cleavage in macrophages. Probing the system with specific cathepsin L, S and B inhibitors would help elucidate the regulatory mechanisms each of the enzymes could be playing in this system.

Cathepsin L largely colocalizes with cystatin B in MDA-MB-231cells, suggesting the cathepsin L is in the cytoplasm [50]. Cathepsins in the cytoplasm of cells has been shown to induce apoptosis through cleavage of Bid and antiapoptotic proteins in multiple cell types [137, 138]. Despite high cytoplasmic concentrations of cathepsin L, MDA-MB-231cells do not undergo apoptosis, suggesting they have bypassed a conserved apoptosis trigger seen in a variety of cancer and non-cancer cell types. This motivates three lines of questioning: 1) Why are MDA-MB-231cells seemingly impervious to cytoplasmic cathepsin induced apoptosis? 2) Are other cancer cells resistant to cytoplasmic cathepsin induced apoptosis? 3) How can this mechanism be exploited or overcome to improve cancer treatment?

Since cytoplasmic cathepsins induce apoptosis through cleavage of substrates, it is possible MDAs evade apoptosis through the inhibition or inactivation of cathepsin L in the cytoplasm. A likely candidate for inhibiting cathepsin L is cystatin B, which has been shown to colocalize with cathepsin L in these cells. Silencing cystatin B and monitoring apoptosis could reveal if this inhibitor is keeping cytoplasmic cathepsin L in check. Alternatively, cathepsin L could be overexpressed with DNA or mRNA, potentially upsetting the balance between protease and inhibitor that is preventing apoptosis.

Whether a tolerance for cytoplasmic cathepsin L is unique to MDA-MB-231 breast cancer cells is a relatively easy question to answer, requiring immunocytochemistry or cell fractionation combined with protein assays to confirm. While the methodology to answer this question is trivial, the implications of multiple cancer cell types evading cathepsin induced apoptosis could open up a new line of attack for cancer therapy.

Identifying the exact mechanism responsible for cancer cells evading apoptosis by cytoplasmic cathepsins could allow its exploitation for cancer treatment. Assuming intracellular cathepsin inhibitors like cystatin B are responsible, these cathepsin inhibitors could paradoxically be viable therapeutic targets. A cytoplasmically-targeted small molecule cathepsin inhibitor-inhibitor would compete with cathepsins binding cystatin B, potentially freeing cytoplasmic cathepsins to induce apoptosis through cleavage of anti-apoptotic proteins. A cathepsin inhibitor-inhibitor would have the potential to provoke unintended extracellular proteolysis or induce apoptosis or lysosomal dysfunction in off-target cells, but it is an intriguing potential treatment option since most non-cancerous cells do not have significant cytoplasmic cathepsins. Even if this theoretical molecule has minimal use as a human cancer therapeutic it could still be incredibly useful for proteolytic research, since it could be used to disable ubiquitous intra and extracellular cathepsin inhibitors, which obfuscate cathepsin dynamics in a variety of tissues. Additionally, conjugating the cathepsin inhibitor-inhibitor with a fluorophore would allow for monitoring of endogenous cathepsin inhibitor levels inside cells and in extracellular fluids in a potentially noninvasive way.

Our lab frequently cites the successes and failures of cathepsin inhibitors to treat various pathologies as motivation for our work in unraveling proteolytic regulatory networks. While I believe this approach is valid, I have often wondered if we could increase our impact on future cathepsin inhibitor treatments by studying the side effects of these inhibitors more directly. Skin lesions are among the most common side effects of cathepsin K inhibitors and dermal fibroblasts express cathepsin K, primarily in their lysosomes to facilitate the degradation of endocytosed collagen fragments [43, 139]. Coincidentally,

cancer associated fibroblasts are increasingly recognized as important drivers of cancer growth and metastasis and have also been shown to produce cathepsins [140]. I think we could provide some valuable insight to the pharmaceutical industry by extending the cathepsin inhibitor treatment experiments we have previously done in macrophages and breast cancer, and expand it to a cell type that is likely instrumental in producing some of the side effects that have kept cathepsin inhibitors out of the clinic.

APPENDIX

A.1. Breast cancer inhibitor computational model full details and parameter optimization methods.

Based on the model diagram in Figure 7-4, we derived a system of ordinary differential equations (ODE) using simple mass action (Eqs. 1 – 5) that describe cathepsin dynamics. This system of equations contains five dynamic variables corresponding to the four cathepsin species and E64, as well as, nine kinetic parameters describing the fluxes within the system. Based on our zymography and Western blot data we aimed to identify values of the kinetic parameters that are in agreement with our experimental findings.

$$\frac{\mathrm{d[Pro]}}{dt} = \mathrm{k_{prod}} - \mathrm{k_{act}} \cdot \mathrm{[Pro]} \tag{1}$$

$$\frac{\mathrm{d[Mat]}}{dt} = \mathrm{k_{act}} \cdot \mathrm{[Pro]} - \mathrm{k_{inact}} \cdot \mathrm{[Mat]} - \mathrm{k_{deg}} \cdot \mathrm{[Mat]} - \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} \tag{2}$$

$$\frac{d[\mathrm{Mat_{inact}}]}{dt} = \mathrm{k_{inact}} \cdot \mathrm{[Mat]} - \mathrm{k_{degI}} \cdot \mathrm{[Mat_{inact}]} \tag{3}$$

$$\frac{\mathrm{d[MatE64]}}{dt} = \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} - \mathrm{k_{degE64}} \cdot \mathrm{[MatE64]} \tag{4}$$

$$\frac{\mathrm{d[E64]}}{dt} = \mathrm{k_{in,E64}} - \mathrm{k_{out,E64}} \cdot \mathrm{[E64]} - \mathrm{k_{E64}} \cdot \mathrm{[Mat]} \cdot \mathrm{[E64]} \tag{5}$$

Parameterization of the dynamic model was performed using a steady state approach. We assume that the E64 concentration in the lysosomal compartment, which is where cathepsin S is localized, reaches a steady state concentration equal to the specified media concentration after 24 hours. Steady state concentrations in the cytosol, which is where cathepsin L is localized, were scaled based on an approximation that the cytosol volume is 35 times larger than the lysosomal volume. Based on these assumptions, and the

proposed system of ODEs (Eqs. 1 – 5), steady state equations were formulated for the three tested concentrations of E64: 0 μM, Eqs. 6 – 8; 10 μM, Eqs. 9 – 12; 50 μM, Eqs. 13 – 16. This system of equations includes seven unknown steady state kinetic parameters, as well as steady state concentrations of cathepsin species that are not directly measurable based on available experimental data, including inactive and inhibitor bound cathepsins.

$$k_{prod} - k_{act} \cdot [Pro]_{0\mu M} = 0 \tag{6}$$

$$k_{act} \cdot [Pro]_{0\mu M} - k_{inact} \cdot [Mat]_{0\mu M} - k_{deg} \cdot [Mat]_{0\mu M} = 0 \tag{7}$$

$$k_{inact} \cdot [Mat]_{0\mu M} - k_{degI} \cdot [Mat_{inact}]_{0\mu M} = 0 \tag{8}$$

$$k_{prod} - k_{act} \cdot [Pro]_{10\mu M} = 0 \tag{9}$$

$$\begin{aligned} k_{act} \cdot [Pro]_{10\mu M} - k_{inact} \cdot [Mat]_{10\mu M} - k_{deg} \cdot [Mat]_{10\mu M} \\ - k_{E64} \cdot [Mat]_{10\mu M} \cdot [E64]_{10\mu M} = 0 \end{aligned} \tag{10}$$

$$k_{inact} \cdot [Mat]_{10\mu M} - k_{degI} \cdot [Mat_{inact}]_{10\mu M} = 0 \tag{11}$$

$$k_{E64} \cdot [Mat]_{10\mu M} \cdot [E64]_{10\mu M} - k_{degE64} \cdot [MatE64]_{10\mu M} = 0 \tag{12}$$

$$k_{prod} - k_{act} \cdot [Pro]_{50\mu M} = 0 \tag{13}$$

$$\begin{aligned} k_{act} \cdot [Pro]_{50\mu M} - k_{inact} \cdot [Mat]_{50\mu M} - k_{deg} \cdot [Mat]_{50\mu M} \\ - k_{E64} \cdot [Mat]_{50\mu M} \cdot [E64]_{50\mu M} = 0 \end{aligned} \tag{14}$$

$$k_{inact} \cdot [Mat]_{50\mu M} - k_{degI} \cdot [Mat_{inact}]_{50\mu M} = 0 \tag{15}$$

$$k_{E64} \cdot [Mat]_{50\mu M} \cdot [E64]_{50\mu M} - k_{degE64} \cdot [MatE64]_{50\mu M} = 0 \quad (16)$$

Eqs. 17 – 22 relate the Western blot and zymography measurements for the specified E64 concentration to the corresponding steady state concentrations of the different species of mature cathepsin. Procathepsin concentrations are absent from these equations, since steady concentrations were determined by the procathepsin bands of the corresponding Western blots. At 0 μM of E64, the mature cathepsin species are computable directly from the Western blot and zymography, while at 10 and 50 μM of E64 the steady state model is required to determine the steady state concentrations of some mature cathepsin species.

$$WB_{0\mu M} = [Mat]_{0\mu M} + [Mat_{inact}]_{0\mu M} \quad (17)$$

$$Zymo_{0\mu M} = [Mat]_{0\mu M} \quad (18)$$

$$WB_{10\mu M} = [Mat]_{10\mu M} + [Mat_{inact}]_{10\mu M} + [MatE64]_{10\mu M} \quad (19)$$

$$Zymo_{10\mu M} = [Mat]_{10\mu M} \quad (20)$$

$$WB_{50\mu M} = [Mat]_{50\mu M} + [Mat_{inact}]_{50\mu M} + [MatE64]_{50\mu M} \quad (21)$$

$$Zymo_{50\mu M} = [Mat]_{50\mu M} \quad (22)$$

The steady state equations (Eqs. 6 – 16) can be simplified based on the expressions relating the experimental measurements with the steady state cathepsin concentrations (Eqs. 17 – 22). Eqs. 23 – 28 define the steady state concentrations of mature cathepsin species at 10 and 50 μM of E64 in terms of experimental measurements and the kinetic parameters.

$$[Mat]_{10\mu M} = Zymo_{10\mu M} \quad (23)$$

$$[\mathrm{Mat_{inact}}]_{10\mu M} = [\mathrm{Mat}]_{10\mu M} \cdot \frac{k_{inact}}{k_{degI}} \tag{24}$$

$$[\mathrm{MatE64}]_{10\mu M} = [\mathrm{Mat}]_{10\mu M} \cdot [\mathrm{E64}]_{10\mu M} \cdot \frac{k_{E64}}{k_{degE64}} \tag{25}$$

$$[\mathrm{Mat}]_{50\mu M} = Zymo_{50\mu M} \tag{26}$$

$$[\mathrm{Mat_{inact}}]_{50\mu M} = [\mathrm{Mat}]_{50\mu M} \cdot \frac{k_{inact}}{k_{degI}} \tag{27}$$

$$[\mathrm{MatE64}]_{50\mu M} = [\mathrm{Mat}]_{50\mu M} \cdot [\mathrm{E64}]_{50\mu M} \cdot \frac{k_{E64}}{k_{degE64}} \tag{28}$$

By using the expressions in equations Eqs. 23 – 28, the steady state equations (Eqs. 6 – 16) were reduced to a system of six equations (Eqs. 29 – 34) with six cathepsin-related kinetic parameters as unknowns, while the cathepsin production rate can later be determined based on Eq. 6. Eqs. 29 – 32 define the relationship between the kinetic parameters and the steady state concentrations, while Eqs. 33 and 34 serve as constraints connecting the steady state concentrations to the experimental data. Additionally, the observation that procathepsin concentrations were not significantly affected by E64 concentration was used to simplify the model further.

$$[\mathrm{Mat}]_{10\mu M} - [\mathrm{Pro}]_{0\mu M} \cdot \frac{k_{act}}{k_{inact} + k_{deg} + k_{E64} \cdot [\mathrm{E64}]_{10\mu M}} = 0 \tag{29}$$

$$[\mathrm{Mat}]_{50\mu M} - [\mathrm{Pro}]_{0\mu M} \cdot \frac{k_{act}}{k_{inact} + k_{deg} + k_{E64} \cdot [\mathrm{E64}]_{50\mu M}} = 0 \tag{30}$$

$$\mathrm{k_{act}} \cdot [\mathrm{Pro}]_{0\mu M} - [\mathrm{Mat}]_{0\mu M} \cdot \left(\mathrm{k_{inact}} + \mathrm{k_{deg}}\right) = 0 \tag{31}$$

$$k_{inact} \cdot [Mat]_{0\mu M} - k_{degI} \cdot [Mat_{inact}]_{0\mu M} = 0 \quad (32)$$

$$[Mat]_{10\mu M} + [Mat_{inact}]_{10\mu M} + [MatE64]_{10\mu M} - WB_{10\mu M} = 0 \quad (33)$$

$$[Mat]_{50\mu M} + [Mat_{inact}]_{50\mu M} + [MatE64]_{50\mu M} - WB_{50\mu M} = 0 \quad (34)$$

Parameter estimation was initially performed for cathepsin S and L based on the Eqs. 23 – 34, where the squared residuals for Eqs. 29 – 34 were minimized with respect to the six kinetic parameters. These equations define a nonlinear root finding problem composed of six equations and six unknowns, where an exact solution cannot be determined analytically. Therefore, a numerical approximate solution was obtained using a newly developed optimization algorithm for black-box problems by identifying kinetic parameter values that minimized the sum of the squared residuals for Eqs. 29 – 34. Based on this model, parameter estimation for cathepsin L proved successful with the identification of a solution with low residuals, but for the case of cathepsin S no reasonable solution was found.

Further analysis lead to redefining the relationship between the zymography measurements and steady state cathepsin concentrations, specifically mature cathepsin S bound to E64 was found to be detectable by zymography. Therefore, for cathepsin S the relationships in Eqs. 20 and 22 were replaced with Eqs. 35 and 36.

$$Zymo_{10\mu M} = [Mat]_{10\mu M} + [MatE64]_{10\mu M} \quad (35)$$

$$Zymo_{50\mu M} = [Mat]_{50\mu M} + [MatE64]_{50\mu M} \quad (36)$$

With Eqs. 35 and 36 the simplification of the steady state parameters is no longer possible, therefore the cathepsin S parameters were identified using the full set of steady state equations, including Eqs. 6 – 19, 21, 35, and 36. Based on these equations, a similar optimization approach using the black-box optimization algorithm was used to determine

the steady state kinetic parameters. For the updated cathepsin S model, there are 13 unknowns including seven kinetic parameters and six steady state concentrations for mature cathepsin S species. The parameter estimation was performed by minimizing the squared residuals for Eqs. 6 – 19, 21, 35, and 36 with respect to all 13 unknowns, but the final steady state concentrations for cathepsin S species were back calculated based on the steady state kinetic parameters.

Dynamic simulations were performed based on the steady state kinetic parameters and the system of ODEs defined by Eqs. 1 – 5. The rate of influx and efflux for E64 were approximated to ensure that the assumed steady state E64 concentrations were reached within 24 hours. Additionally, the kinetic parameters were scaled to ensure the system reached a steady state within 24 hrs. Numerical simulations were performed using the deSolve package in the R statistical language.

A.2. Cathepsin S and K overexpression model parameter estimation details

A total of 12 model parameters were estimated using COPASI's built-in genetic algorithm implemented on Georgia Tech's Partnership for an Advanced Computing Environment (PACE) Cluster. The genetic algorithm was implemented 100 times, with randomized initial conditions, population of 10,000 and run for 200 generations, resulting in approximately $2x10^6$ function evaluations per run. Default COPASI genetic algorithm settings were used (Log verbosity = 0, Random number generator = 1, Seed = 0, Mutation variance = 0.1). Parameter space was bounded as indicated in Table A.2-1. Upper bounds for reaction on rates were restricted to below the measured binding rate of streptavidin on biotin, one of the strongest naturally occurring, non-covalent molecular interactions currently known [141]. Bounds for reaction off rates and catalytic rates were set based on experimentally measured extremes of these rates previously documented in different enzymes [142]. Secretion rates were initially allowed to vary between $1x10^{-6}$ and $1x10^6$, but repeated runs consistently estimated the secretion rates to fall between 1 and 200. General degradation rates were initially bounded between $1x10^{-6}$ and $1x10^6$, but solutions with high general degradation rates had very little impact of cathepsins on cathepsin degradation, which does not agree with the experimental data. Degradation rates were capped to force more of the solutions to involve cathepsins in cathepsin degradation.

Table A.2 - 1: Lower and upper bounds for parameter estimation

Parameter	Lower Bound	Upper Bound
Ssec	1	200
Sact	$1x10^{-6}$	$1x10^{6}$
Sdeg	$1x10^{-6}$	10
Ksec	1	200
Kact	$1x10^{-6}$	$1x10^{6}$
Kdeg	$1x10^{-6}$	10
SonK	$1x10^{-6}$	3.6
SoffK	$1x10^{-4}$	$1x10^{6}$
ScatK	$1x10^{-2}$	$1x10^{7}$
KonK	$1x10^{-6}$	3.6
KoffK	$1x10^{-4}$	$1x10^{6}$
KcatK	$1x10^{-2}$	$1x10^{7}$

A.3. Identification of Critical Quality Attributes in Mesenchymal Stem Cells

A.3.1. Introduction

Georgia Tech, University of Wisconsin-Madison, University of Georgia and University of Puerto Rico Mayaguez and other academic and industry partners have been awarded a National Science Foundation Engineering Research Center for the purpose of developing Cell Manufacturing Technologies (CMaT). The goal of the center is to "enable robust, scalable, low-cost biomanufacturing of high-quality therapeutic cells to bring affordable, curative therapies against incurable chronic diseases". Cell culture conditions are a key component of cell manufacturing and identifying non-invasive indicators of optimal cell culture conditions and high-quality cells will be instrumental to increasing cell manufacturing output while keeping quality high and lowering costs. A major initiative of CMaT is identifying critical quality attributes (CQAs), capable of being monitored non-invasively during cell manufacture, for the production of Chimeric Antigen Receptor (CAR) T cells for cancer immunotherapy, induced pluripotent stem cells (iPSC) cardiomyocytes for cardiac regeneration and mesenchymal stem cells (MSCs) for immune modulation and musculoskeletal repair. Once CQAs have been identified, robust cell manufacturing processes can be designed, capable of monitoring cell quality at multiple stages during development.

This appendix focuses on identifying indicators of poor cell culture conditions in MSCs. MSCs are multipotent progenitor cells found in bone marrow, adipose tissue, amniotic membranes and umbilical cords capable of differentiating into osteoblasts, adipocytes and chondrocytes [143]. MSCs derived from adult bone marrow are currently the most popular tissue source [144]. MSCs, particularly those isolated from adult bone marrow, display significant inter-donor variability in gene expression, morphology and cytokine secretion [143, 145]. Due to this high degree of donor variability, this pilot study

was conducted to identify cytokines secreted by CMaT MSCs and observe the effects of suboptimal cell culture conditions on the MSC secretory profiles.

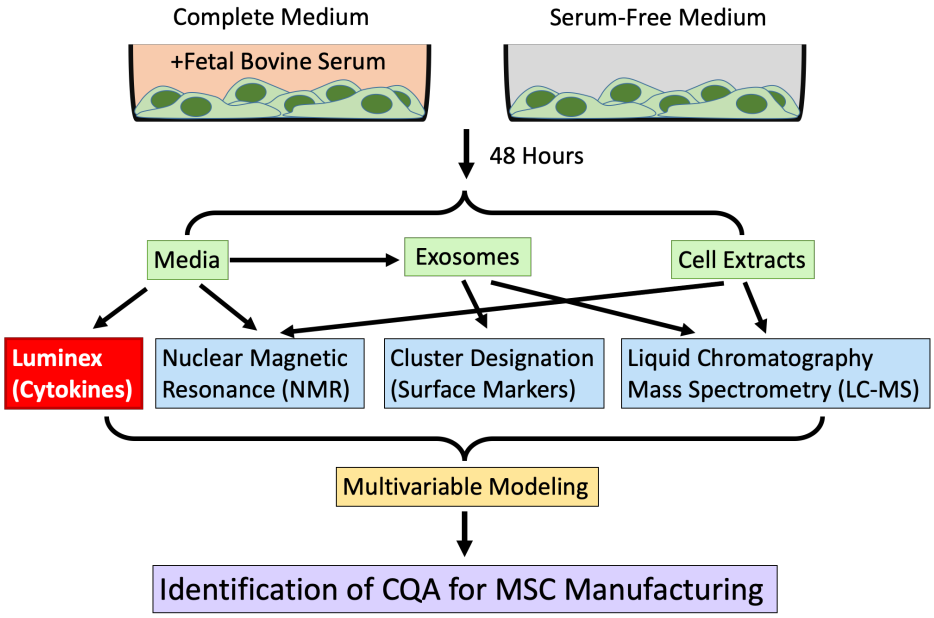

Figure A.3 - 1: Experimental design to identify indicators of poor cell culture conditions in mesenchymal stem cells (MSCs).
MSCs were grown in media supplemented with 10% fetal bovine serum or in media containing no fetal bovine serum for 48 hours. Conditioned media, exosomes and cell extracts were then analyzed for proteins and metabolites by multiple research groups using Luminex (detailed here), Nuclear Magnetic Resonance, Cluster Designation and Liquid Chromatography Mass Spectrometry. Data collected from these techniques will then be used to develop a multivariable model capable of identifying critical quality attributes for industrial MSC culture.

A.3.2. Materials and Methods

A.3.2.1. Mesenchymal Stem Cell Culture

Mesenchymal stem cells, induced from pluripotent stem cells (Cellular Dynamics lot 0003) were plated in 150mm plates at a density of 500 cells/cm2 and cultured at 37°C and 5% CO2 in MEM-α media (Gibco) supplemented with 1% L-Glutamine 200mM (Gibco), 1% Penicillin Streptomycin (Gibco) and with or without 10% Fetal Bovine Serum

(Hyclone). Cells were passaged upon reaching 80% confluency. Media was collected after 48 hours of culture in complete or serum-free media.

A.3.2.2. Human Luminex Performance Assay

Human magnetic Luminex performance assay kit and a customized panel of cytokine microparticles were purchased from R&D Systems and assay was carried out according to the manufacturer's instructions. Briefly, conditioned media samples, either complete or serum-free, were incubated with microparticles for 3 hours. Biotin conjugated antibodies were added to the microparticles and incubated for 1 hour, followed by Streptavidin-PE and another 30 minutes of incubation. Fluorescence of the microparticles was measured in a Bio-Rad Bio-Plex 200 System. All incubations were done at room temperature on a shaker at 800 rpm. All wash steps were done using a magnetic plate.

Table A.3 - 1: Cytokines profiled in MSC conditioned media

Analyte	Function
CCL2	Recruits immune cells to sites of inflammation
Angiopoeitin-1	Promote angiogenesis and vascular growth
IL-6	Numerous, pro-inflammatory signaling
Cystatin C	Inhibitor of cysteine cathepsin proteases
IL-8/CXCL8	Stimulates neutrophil migration and phagocytosis
VEGF	Encourages formation of blood vessels,
TNF-alpha	Regulates immune cells and systemic inflammation
CCL3/MIP-1alpha	Recruits and activates granulocytes to sites of inflammation
FGF basic	Numerous, cell growth and development, embryonic development
G-CSF	Stimulates granulocyte and stem cell production and release
Eotaxin	Recruits and activates Eosinophils to sites of inflammation

EGF	Promotes cellular proliferation, differentiation and survival

A.3.3. Results

Mesenchymal stem cells were cultured with complete media (containing fetal bovine serum) or in serum-free media for 48 hours, after which cell extracts, conditioned media and media exosomes were collected for analysis with a variety of analytical methods (Figure A.3-1). Conditioned media and cell-free media controls were probed for 12 different cytokines, which have previously been identified in MSC media (Table A.3-1).

Out of the 12 cytokines measured, 10 were detectable in media samples and 6 were significantly different between MSC conditioned complete and serum-free media (Figure A.3-2). CCL2/MCP-1, Angiopoietin-1, IL-6, Cystatin C and IL-8/CXCL8 were significantly lower in serum-free than complete media, while VEGF was significantly higher in serum-free media than complete media. TNFα, CCL3/MIP-1α, FGF-Basic and G-CSF were not significantly different between the MSC complete and serum-free conditions. Most of the cytokines measured were nearly undetectable in cell-free media controls, indicating the cytokines are produced by the MSCs and not found in fetal bovine serum, with the exception of angiopoietin-1.

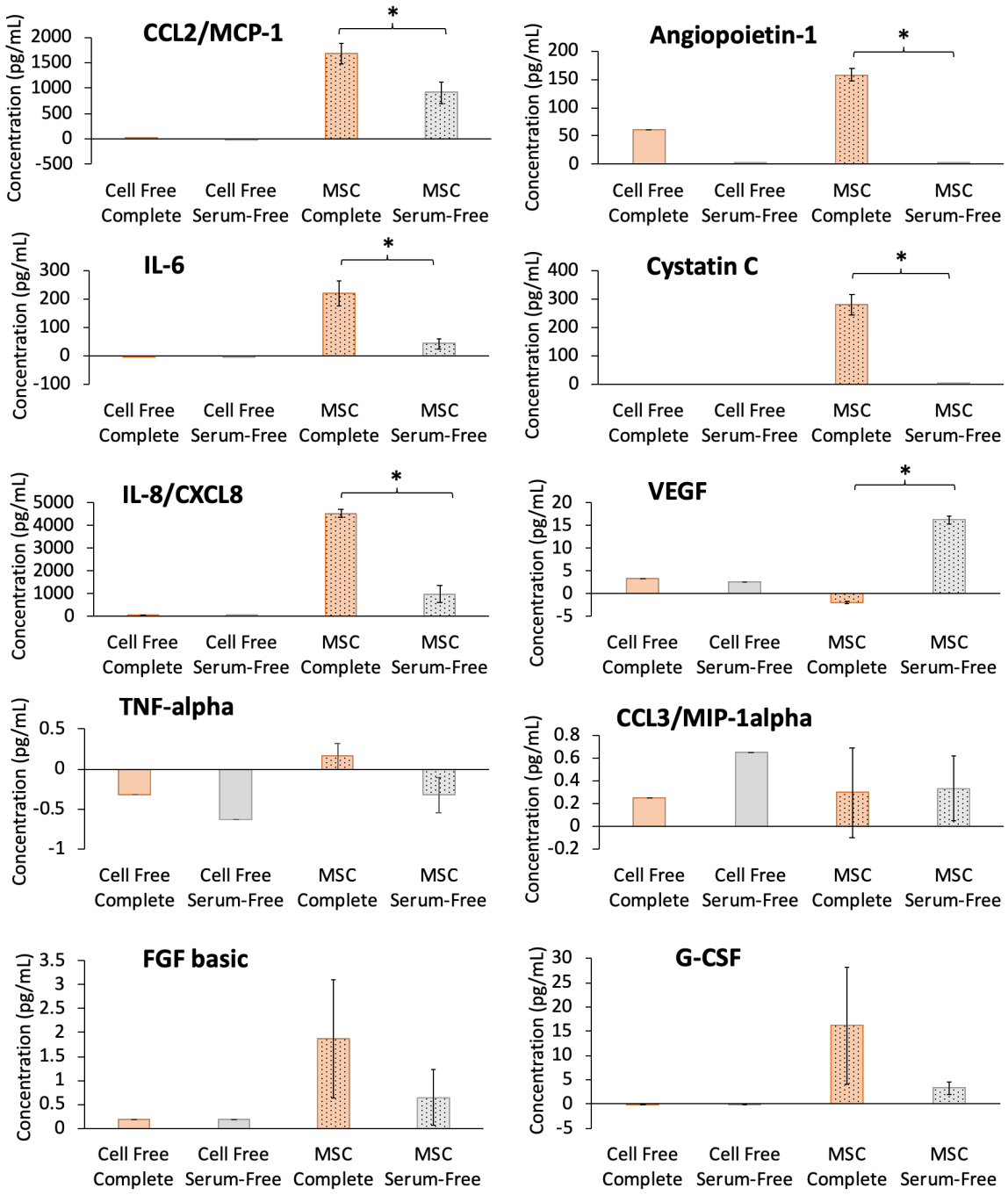

Figure A.3 - 2: Serum affects cytokines found in MSC media.
CCL2/MCP-1, Angiopoietin-1, IL-6, Cystatin C, IL-8/CXCL8 were found in significantly greater concentrations in conditioned complete media than serum-free media. VEGF was significantly greater in serum-free than complete media. TNF-alpha, CCL3/MIP-1alpha, FGF basic and G-CSF were not different between complete and serum-free media. Angiopoietin-1 and VEGF were detectable in small amounts in cell-free complete media,

but no other analytes were found in cell-free media. Eotaxin and EGF were also measured, but were undetectable in any samples. n=3 *p<0.05 by t-test.

A.3.4. Discussion

This study provides proof of concept that cytokine measurements can be used to differentiate between mesenchymal stem cells grown under different culturing conditions. Cytokine profiling of mesenchymal stem cell media is a non-invasive method of monitoring cell health that could be implemented in larger scale manufacturing processes to continuously measure the quality of new cells.

The Luminex results show a variety of cytokines are detectable in MSC media and that many of them are responsive to media serum content. Importantly, most of the cytokines were extremely rare in cell-free media, indicating these cytokines are being produced by the MSCs and not supplied by the media or serum. Several of the cytokines that were lower in serum-free media function primarily to recruit immune cells (CCL2/MCP-1, IL-6, IL-8/CXCL8) indicating loss of serum could interfere with immune cell recruitment by MSCs. Interestingly, we observed opposite responses in two angiogenic proteins with Angiopoietin-1 being lower in serum-free media while VEGF was increased in serum-free media, indicating serum affecting blood vessel growth in a complicated manner. Cystatin C, the endogenous protein inhibitor of cysteine cathepsins secreted constitutively by most human cell types, was almost undetectable in MSC serum-free media. Transcripts of multiple cystatins, including cystatin C, are significantly lower in human fetal liver derived MSCs compared to human adult bone marrow derived MSCs, showing cystatin C secretion varies among different MSC sources [146]. Secretion of cystatin C decreases upon differentiation of human bone marrow derived MSCs to osteoblasts, suggesting cystatin C could be a useful biomarker for MSC potency [147].

This data can also be used to construct multivariable partial least square regression models capable of predicting cell health and cell culture conditions based on the secreted

cytokine profile detectable in cell culture media. Future studies will examine the differences in cytokine secretomes of MSCs from different sources, including bone marrow and induced pluripotent stem cell derived MSCs. Identifying cytokines with conserved responses to cell culture conditions across different MSC lineages will ensure model predictions are robust and applicable to the many sources of MSCs currently under investigation for therapeutic applications.

A.4. Aggressive Breast Cancer in Ethiopian Women

A.4.1. Introduction

In the United States, median age of breast cancer diagnosis is 63 years for white women and 59 years for black women, and about half of breast cancer is diagnosed while still localized at the primary site [116]. At Black Lion Hospital in Addis Ababa, Ethiopia (one of the largest hospitals in the country) the median age of breast cancer diagnosis is 40 years and a majority of cases have already metastasized out of the primary tumor site by time of diagnosis. Through our collaboration with the researchers at Black Lion Hospital and Addis Ababa University, we hope to assist in the discovery of the mechanisms responsible for this extremely aggressive cancer seen in Ethiopian women. One potential hypothesis we are exploring is upregulation of active proteases, including cysteine cathepsins, in aggressive breast cancer. We also investigated the abundance of the inflammatory cytokines TNFα and IL-6, as well as cystatin C, in plasma from patients undergoing chemotherapy at Black Lion Hospital.

A.4.2. Materials and Methods

A.4.2.1. Preparation of human samples

Primary tumor and patient-matched non tumor mammary tissue were collected from six patients during mastectomy and frozen. Samples were homogenized by grinding in a mortar and pestle with zymography lysis buffer. Whole blood from 26 patients before

and after chemotherapy was collected in heparinized tubes, centrifuged to remove red and white blood cells and frozen.

A.4.2.2. Cathepsin zymography

Protein concentration of homogenized samples were determined by BCA and equal amounts of protein were analyzed by cathepsin zymography as previously described [84, 85].

A.4.2.3. Human Luminex Performance Assay

Human magnetic Luminex performance assay kit and a customized panel of cytokine microparticles were purchased from R&D Systems and assay was carried out according to the manufacturer's instructions. Briefly, equal volumes of patient serum were incubated with microparticles for the analytes of interest. Biotin conjugated antibodies were added to the microparticles and incubated for 1 hour, followed by Streptavidin-PE and another 30 minutes of incubation. Fluorescence of the microparticles was measured in a Bio-Rad Bio-Plex 200 System. All incubations were done at room temperature on a shaker at 800 rpm. All wash steps were done using a magnetic plate.

A.4.3. Results

A.4.3.1. Active cathepsins overexpressed in Ethiopian breast cancer tissue samples

Equal amounts of homogenized primary cancer (C) and patient-matched normal tissue (N) from six patients were loaded and analyzed by cathepsin zymography. Active cathepsins were clearly detectable in cancer tissue from most of the patients, primary a 20 kDa band consistent with active cathepsin L. An approximately 35 kDa band was also detectable in cancer tissue from one patient, potentially caused by cathepsin S or cathepsin V. These results support the hypothesis that active cathepsins are involved in Ethiopian breast

cancer. Additional experiments will be required to compare the relative active proteases in samples from Ethiopian patients to other patient groups.

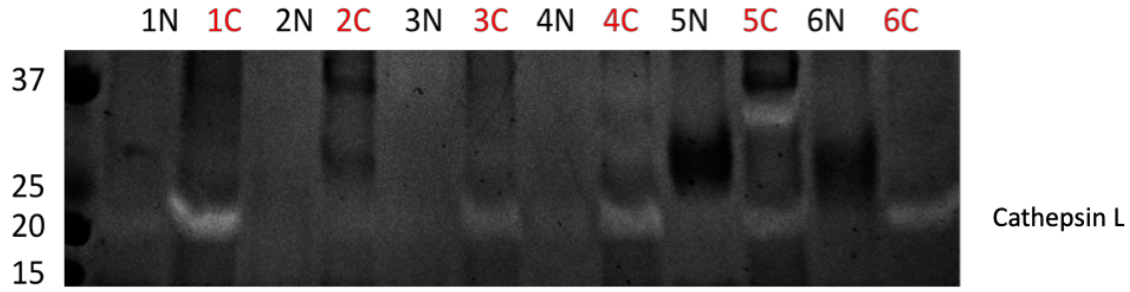

Figure A.4 - 1: Active cathepsins are elevated in breast cancer tissue compared to patient-matched controls.
Cancer (C) and normal (N) tissue from six Ethiopian breast cancer patients was homogenized in a mortar and pestle and equal amounts of protein were loaded for cathepsin zymography. Active cathepsin L was stronger in the cancer than normal tissue and active cathepsin S was detectable in one cancer sample and none of the normal samples.

A.4.3.2. Cytokine expression in Ethiopian breast cancer patient plasma

Levels of TNFα, IL-6 and cystatin C in serum from Ethiopian breast cancer patients undergoing chemotherapy at Black Lion Hospital were measured by Luminex. The average values of each cytokine are reported in Table A.4-1. No statistically significant differences in cytokine expression were observed following chemotherapy. Cystatin C expression was found to be approximately 100,000-fold higher than TNFα or IL-6 levels.

Table A.4-1: Average cytokine levels in plasma from Ethiopian breast cancer patients

(pg/mL)	Pre-Chemotherapy		Post-Chemotherapy	
TNFα	4.2	± 2.3	4.7	± 1.8
IL-6	6.4	± 11.4	2.9	± 4.3
Cystatin C	283962.1	± 143066.8	310618.7	± 133230.8

When cytokine levels in individual patients were compared pre and post chemotherapy, a significant increase in TNFα was observed, while IL-6 and Cystatin C were not significantly different (Figure A.3-2A&B). The majority of patients had extremely low or undetectable levels of IL-6 (Figure A.3-2A). However, among patients with initial IL-6 levels above 10 pg/mL, IL-6 was significantly decreased following chemotherapy. Cystatin C levels were not significantly different before and after chemotherapy, but there was some variability in cystatin C levels among the patients, ranging from 57,000 to 880,000 pg/mL with most samples falling close to the mean value of 280,000 pg/mL.

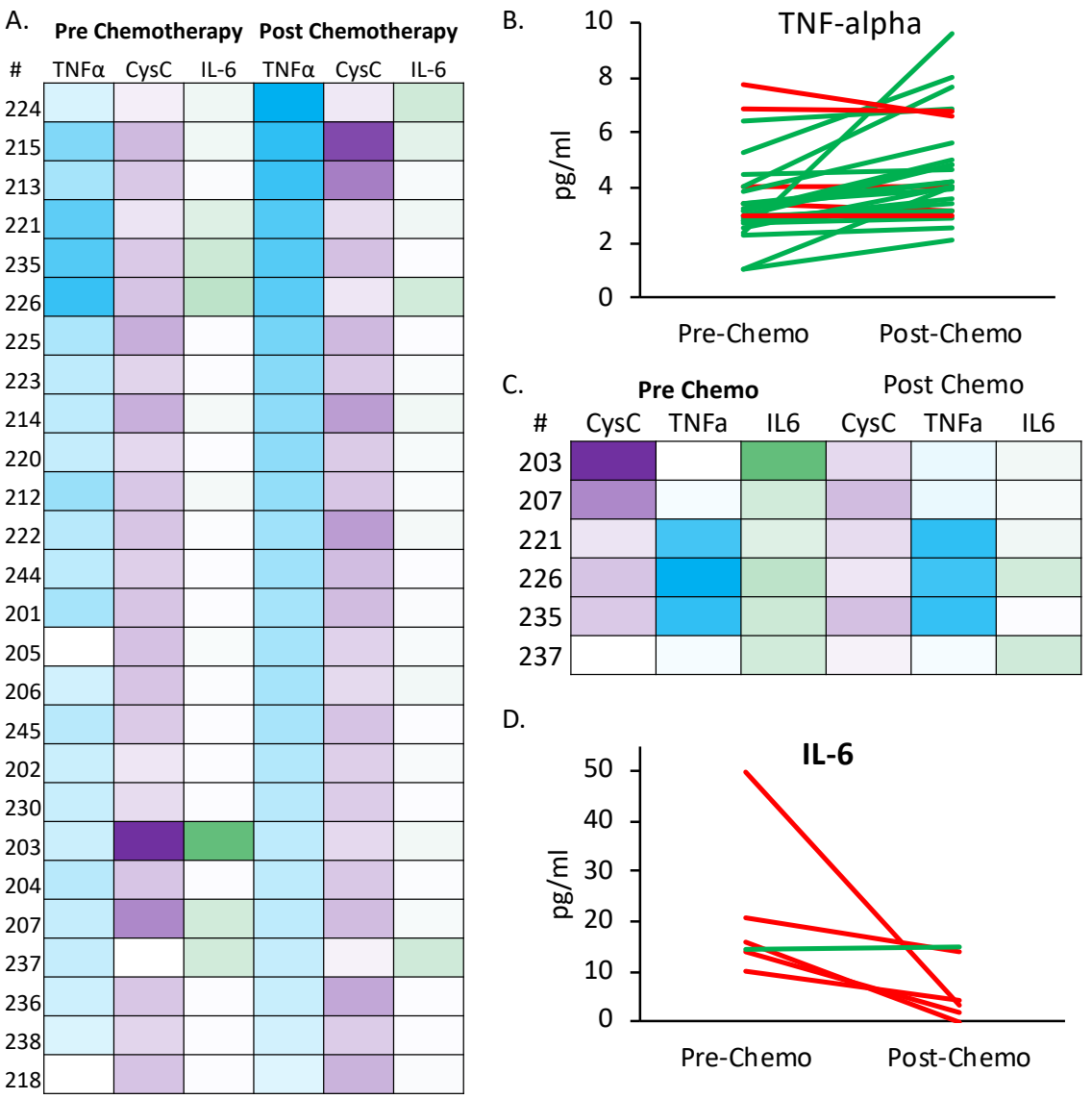

Figure A.4 - 2: Plasma TNFα increases and IL-6 decreases in Ethiopian breast cancer patients following chemotherapy.
Plasma TNFα, IL-6 and cystatin C levels before and after chemotherapy were measured using a multiplex magnetic bead Luminex assay and displayed in a heat map, where darker colors indicate a higher relative amount of the target protein in the plasma samples (A). Plasma TNFα increased following chemotherapy in most of the subjects (n=26 p=0.0014 by matched pairs t-test). Each line represents a patient, green lines have a positive slope, indicating an increase in TNFα while red lines have a negative slope, indicating a decrease (B). IL-6 was undetectable in most of the patient samples, but among patients with high initial plasma IL-6, IL-6 decreased following chemotherapy shown in a heat map (C) and line plot (D) (n=6 p=0.044 by matched pairs t-test).

A.4.4. Discussion

Active cathepsins were elevated in primary tumors compared to patient matched normal tissue. This supports the hypothesis that active cathepsins play a role in breast cancer in Ethiopian women. Levels of TNFα, IL-6 and cystatin C in the plasma of Ethiopian cancer patients were quantified before and after chemotherapy. TNFα levels were found to decrease post-chemotherapy in a statistically significant number of patients, and IL-6 levels were found to significantly increase post-chemotherapy among patients with high initial IL-6 levels.

The TNFα levels we observed in the samples from Ethiopian patients was higher than previously reported in a study of 40 Taiwanese women with breast cancer, which measured an average of 1.47 pg/mL by ELISA compared to our averages of 4.2 and 4.7 pg/mL before and after chemotherapy [148]. The IL-6 levels measured here are lower than those previously reported in studies of breast cancer in Egyptian women, which detected an average of 111.38pg/mL across 40 patients, compared to our averages of 11.4 and 4.3 pg/mL before and after chemotherapy [149]. These results suggest some possible avenues of future research to hopefully identify the causes of this aggressive cancer observed in Ethiopian women and development of treatments capable of combating it.

A.5. CM Papain and concentration methods effects on cathepsin cannibalism

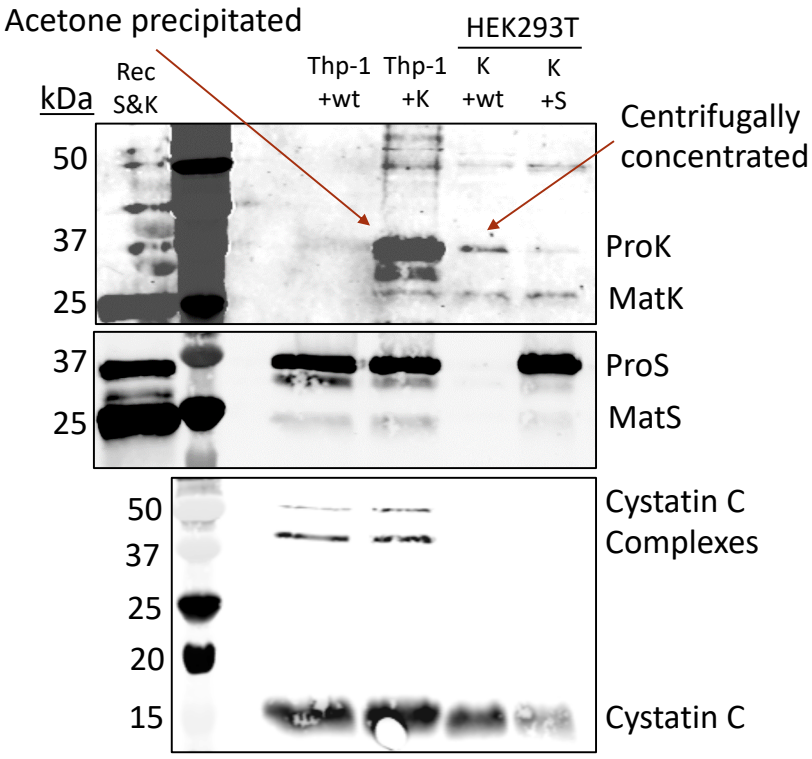

Figure A.5 - 1: Cathepsin K levels are higher in acetone precipitated media from macrophages than in centrifugally concentrated media from HEK293T cells.
Cathepsin and cystatin C levels in concentrated conditioned media from different cocultures were assed by Western blot. Procathepsin K levels were higher in the Thp-1 macrophage cocultures than the HEK293T only cocultures. Cathepsin S levels were similar between the cocultures. Cystatin C levels were slightly higher in macrophage cocultures than HEK293T cocultures. Thp-1 cocultures were concentrated with acetone precipitation while HEK293T only cocultures were concentrated with centrifugal concentrators.

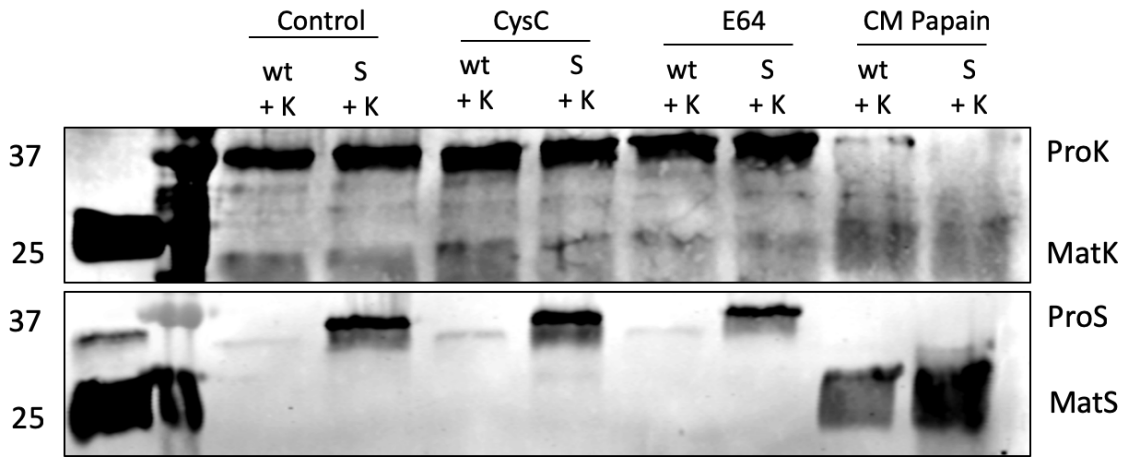

Figure A.5 - 2: Acetone precipitation prevents detection of cannibalism in HEK293T cells and CM Papain appears to allow cathepsin K to be degraded.
HEK293T cells were cocultured with cystatin C mRNA transfection, 50 μM E64 treatment or 1 μM carboxymethyl (CM) papain. No CatS on CatK cannibalism was observed in the control groups when using acetone precipitation. Cystatin C transfection and E64 treatment had no effect on cathepsin K levels. CM papain, a partially deactivated version of papain was used to compete for cystatin C in the media. CM papain treatment reduced procathepsin K levels and increased mature cathepsin S levels. However, CM papain also appeared to be detectable by the cathepsin S antibody used, since a 25 kDa band was detected in the cathepsin S Western blot.

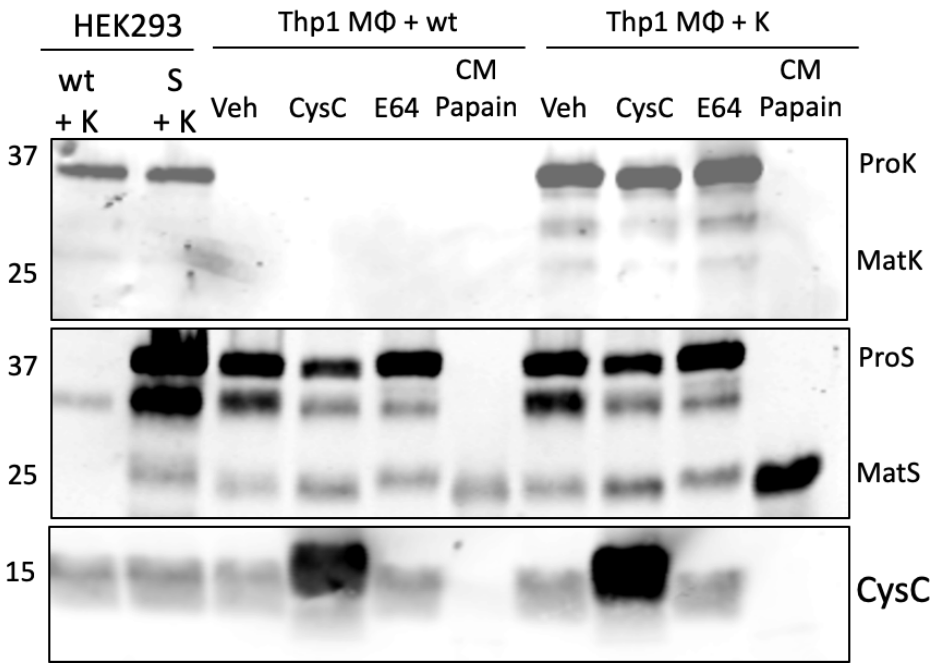

Figure A.5 - 3: CM Papain lowers cathepsin K and cystatin C levels, while increasing mature cathepsin S.
Macrophage cocultures were transfected with cystatin C mRNA or treated with 50 μM E64 or 1 μM CM Papain. Cathepsin K levels were higher in macrophage cocultures than HEK293 cocultures. CM papain abolished pro- and mature cathepsin K, procathepsin S and cystatin C levels. CM papain should not directly interact with cathepsins, but it should bind and not degrade cystatin C. These results suggest there is significant residual activity in the CM papain, likely caused by incomplete methylation of the enzyme. The CM papain appears to have cleaved all of the cystatin C, and potentially could have cleaved cathepsins K and S. Mature cathepsin S appears to be increased following CM papain treatment.

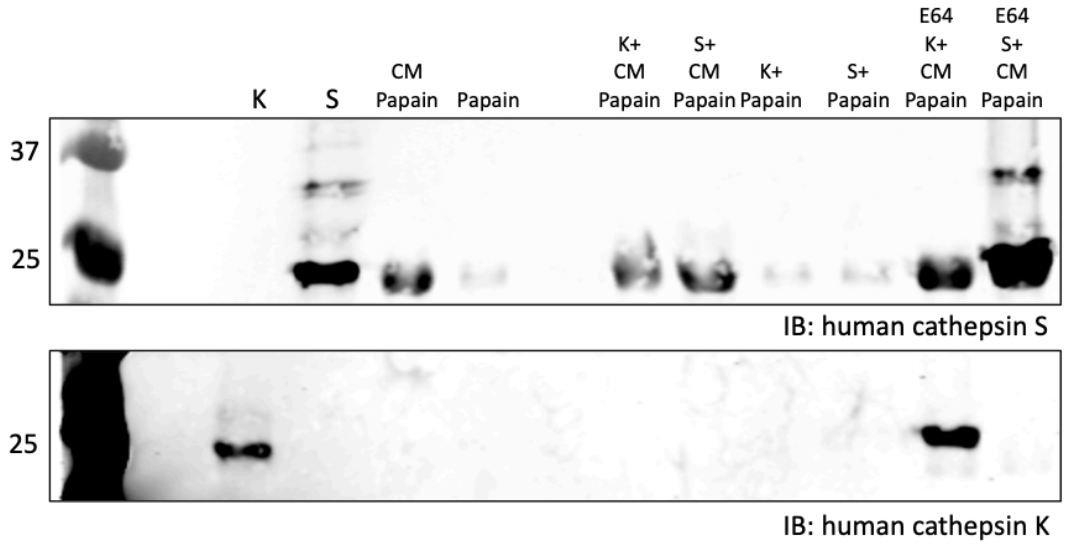

Figure A.5 - 4: CM Papain degrades cathepsins S and K and is detectable by cathepsin S antibody.
100 nM recombinant human cathepsins S and K were incubated in serum-free DMEM for 1 hour at 37°C ±10 µM CM papain or normal papain with ±1 mM E64 then analyzed by Western blot. CM papain and normal papain were both detected by the cathepsin S antibody (polyclonal goat IgG R&D Systems) but not the cathepsin K antibody (Polyclonal rabbit IgG Proteintech). CM papain and normal papain degraded pro- and mature cathepsin S and mature cathepsin K. Degradation was stopped by incubating the enzymes with E64. CM papain and normal papain also appeared to partially auto-degrade during the incubation, since the 25 kDa bands on the cathepsin S blot were darker when incubated with E64. Cathepsins S and K appear to have slightly auto-degraded during the incubation, with E64 treated bands appearing darker than bands for the enzyme incubated alone. Total reaction volumes were 10 µL.

REFERENCES

1. Desantis, C., et al., *Breast cancer statistics , 2013.* CA Cancer Journal for Clinicians, 2014. **64**: p. 52-62.
2. Weigelt, B., J.L. Peterse, and L.J. van 't Veer, *Breast cancer metastasis: markers and models.* Nature reviews. Cancer, 2005. **5**(August): p. 591-602.
3. Scully, O.J., et al., *Breast cancer metastasis.* Cancer genomics & proteomics, 2012. **9**(5): p. 311-20.
4. Mao, Y., et al., *Stromal cells in tumor microenvironment and breast cancer.* Cancer and Metastasis Reviews, 2013. **32**(1): p. 303-315.
5. Kennecke, H., et al., *Metastatic behavior of breast cancer subtypes.* Journal of Clinical Oncology, 2010. **28**(20): p. 3271-3277.
6. Polyak, K., *Heterogeneity in breast cancer.* The Journal of Clinical Investigation, 2011. **121**(10): p. 3786-3788.
7. Williams, C.B., E.S. Yeh, and A.C. Soloff, *Tumor-associated macrophages: unwitting accomplices in breast cancer malignancy.* Nature Publishing Group, 2016(September 2015).
8. Gocheva, V., et al., *IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion.* Genes and Development, 2010. **24**(3): p. 241-255.
9. Chanmee, T., et al., *Tumor-associated macrophages as major players in the tumor microenvironment.* Cancers, 2014. **6**(3): p. 1670-1690.
10. Mantovani, A., et al., *Role of tumor-associated macrophages in tumor progression and invasion*. 2006. p. 315-322.
11. Noy, R. and J.W. Pollard, *Tumor-associated macrophages: From mechanisms to therapy.* Immunity, 2014. **41**(1): p. 49-61.
12. Solinas, G., et al., *Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation.* Journal of leukocyte biology, 2009. **86**(November): p. 1065-1073.
13. Ding, H., et al., *Tumor-associated macrophages induce lymphangiogenesis in cervical cancer via interaction with tumor cells.* Apmis, 2014. **122**(11): p. 1059-1069.
14. Kurahara, H., et al., *M2-polarized tumor-associated macrophage infiltration of regional lymph nodes is associated with nodal lymphangiogenesis and occult nodal involvement in pN0 pancreatic cancer.* Pancreas, 2013. **42**(1): p. 155-9.
15. Lah, T.T., et al., *Cells producing cathepsins D, B, and L in human breast carcinoma and their association with prognosis.* Human pathology, 2000. **31**(2): p. 149-60.
16. Brix, K., et al., *Cysteine cathepsins: Cellular roadmap to different functions.* Biochimie, 2008. **90**(2): p. 194-207.
17. Shi, G.P., et al., *Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease.* The Journal of biological chemistry, 1992. **267**(11): p. 7258-7262.
18. Yasuda, Y., et al., *Cathepsin V, a novel and potent elastolytic activity expressed in activated macrophages.* Journal of Biological Chemistry, 2004. **279**(35): p. 36761-36770.

19. Novinec, M. and B. Lenarčič, *Cathepsin K: a unique collagenolytic cysteine peptidase.* Biological Chemistry, 2013. **394**(9): p. 1163-1179.
20. Platt, M.O. and W.A. Shockey, *Endothelial cells and cathepsins: Biochemical and biomechanical regulation.* Biochimie, 2016. **122**: p. 314-323.
21. Brömme, D., et al., *Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization.* Biochemistry, 1999. **38**(8): p. 2377-2385.
22. McQueney, M.S., et al., *Autocatalytic activation of human cathepsin K.* Journal of Biological Chemistry, 1997. **272**(21): p. 13955-13960.
23. Mehanna, S., et al., *Cathepsin D in pancreatic acinar cells is implicated in cathepsin B and L degradation, but not in autophagic activity.* Biochemical and Biophysical Research Communications, 2015: p. 1-7.
24. Fortelny, N., et al., *Network analyses reveal pervasive functional regulation between proteases in the human protease web.* PLoS Biology, 2014. **12**(5).
25. Mohamed, M.M. and B.F. Sloane, *Cysteine cathepsins: multifunctional enzymes in cancer.* Nature Reviews Cancer, 2006. **6**(10): p. 764-775.
26. Olson, O.C. and J.A. Joyce, *Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response.* Nature Reviews Cancer, 2015. **15**(12): p. 712-729.
27. Castro-Gomes, T., et al., *Plasma-membrane repair is regulated extracellularly by proteases released from lysosomes.* The FASEB Journal, 2015. **29**(1 Supplement): p. 884.42-884.42.
28. Kågedal, K., et al., *Sphingosine-induced apoptosis is dependent on lysosomal proteases.* Biochemical Journal, 2001. **359**(Pt 2): p. 335-343.
29. Vasiljeva, O., et al., *Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans.* FEBS Letters, 2005. **579**(5): p. 1285-1290.
30. Mason, S.D. and J.A. Joyce, *Proteolytic networks in cancer.* Trends in Cell Biology, 2011. **21**(4): p. 228-237.
31. Littlewood-Evans, A.J., et al., *The osteoclast-associated protease cathepsin K is expressed in human breast carcinoma.* Cancer Res, 1997. **57**(23): p. 5386-5390.
32. Jensen, A.B., et al., *The cathepsin k inhibitor odanacatib suppresses bone resorption in women with breast cancer and established bone metastases: Results of a 4-week, double-blind, randomized, controlled trial.* Clinical Breast Cancer, 2010. **10**(6): p. 452-458.
33. Sioud, M. and M.H. Hansen, *Profiling the immune response in patients with breast cancer by phage-displayed cDNA libraries.* Eur J Immunol, 2001. **31**(3): p. 716-725.
34. Sudhan, D.R. and D.W. Siemann, *Cathepsin L targeting in cancer treatment.* Pharmacology & therapeutics, 2015. **155**: p. 105-116.
35. Boudreau, F., et al., *Loss of cathepsin L activity promotes claudin-1 overexpression and intestinal neoplasia.* The FASEB Journal, 2007. **21**(14): p. 3853-3865.
36. Dennemarker, J., et al., *Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis.* Oncogene, 2010. **29**(11): p. 1611-1621.

37. Poole, A.R., et al., *Differences in secretion of the proteinase cathepsin B at the edges of human breast carcinomas and fibroadenomas.* Nature, 1978. **273**(5663): p. 545-547.
38. Ruan, H., et al., *Targeting cathepsin B for cancer therapies.* Horizons in cancer research, 2015. **56**: p. 23-40.
39. Withana, N.P., et al., *Cathepsin B inhibition limits bone metastasis in breast cancer.* Cancer Research, 2012. **72**(5): p. 1199-1209.
40. Shree, T., et al., *Macrophages and cathepsin proteases blunt chemotherapeutic response in breast cancer.* Genes and Development, 2011. **25**(23): p. 2465-2479.
41. Mullard, A., *Merck & Co. drops osteoporosis drug odanacatib.* Nature Reviews Drug Discovery, 2016. **15**(10): p. 669-669.
42. Palermo, C. and J.A. Joyce, *Cysteine cathepsin proteases as pharmacological targets in cancer.* Trends in Pharmacological Sciences, 2008. **29**(1): p. 22-28.
43. Drake, M.T., et al., *Cathepsin K inhibitors for osteoporosis: Biology, potential clinical utility, and lessons learned.* Endocrine Reviews, 2017(June 2017): p. 325-350.
44. Kramer, L., D. Turk, and B. Turk, *The future of cysteine cathepsins in disease management.* Trends Pharmacol Sci, 2017. **38**(10): p. 873-898.
45. Duong, L.T., *Therapeutic inhibition of cathepsin K—reducing bone resorption while maintaining bone formation.* BoneKEy Reports, 2012. **1**(5): p. 1-8.
46. Garber, K., *Two pioneering osteoporosis drugs finally approach approval.* Nature Reviews Drug Discovery, 2016. **15**(7): p. 445-446.
47. Chen, J.C., et al., *Design and synthesis of α-ketoamides as cathepsin s inhibitors with potential applications against tumor invasion and angiogenesis.* Journal of Medicinal Chemistry, 2010.
48. Wijkmans, J. and J. Gossen, *Inhibitors of cathepsin K: a patent review (2004 – 2010).* Expert Opinion on Therapeutic Patents, 2011.
49. Payne, C.D., et al., *Pharmacokinetics and pharmacodynamics of the cathepsin S inhibitor, LY3000328, in healthy subjects.* British Journal of Clinical Pharmacology, 2014. **78**(6): p. 1334-1342.
50. Wilder, C.L., et al., *Differential cathepsin responses to inhibitor-induced feedback: E-64 and cystatin C elevate active cathepsin S and suppress active cathepsin L in breast cancer cells.* The International Journal of Biochemistry & Cell Biology, 2016. **79**: p. 199-208.
51. Erickson, A.H., *Biosynthesis of lysosomal endopeptidases.* Journal of Cellular Biochemistry, 1989(40): p. 31-41.
52. Collette, J., et al., *Biosynthesis and alternate targeting of the lysosomal cysteine protease cathepsin L.* International Review of Cytology, 2004. **241**(SPEC.ISS.): p. 1-51.
53. Reiser, J., B. Adair, and T. Reinheckel, *Specialized roles for cysteine cathepsins in health and disease.* The Journal of Clinical Investigation, 2010. **120**(10): p. 3421-3431.
54. Ludwig, T., et al., *Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts.* The EMBO Journal, 1994. **13**(15): p. 3430-3437.

55. Gabel, C.A., D.E. Goldberg, and S. Kornfeld, *Identification and characterization of cells deficient in the mannose 6-phosphate receptor: Evidence for an alternate pathway for lysosomal enzyme targeting.* Cell Biology, 1983. **80**(February): p. 775-779.
56. Schnyder, J. and M. Baggliolini, *Secretion of lysosomal hydrolases by stimulated and nonstimulated macrophages.* The Journal of experimental medicine, 1978.
57. Hakala, J.K., et al., *Lysosomal enzymes are released from cultured human macrophages, hydrolyze LDL in vitro, and are present extracellularly in human atherosclerotic lesions.* Arteriosclerosis, Thrombosis, and Vascular Biology, 2003. **23**(8): p. 1430-1436.
58. Riese, R.J., et al., *Cathepsin S activity regulates antigen presentation and immunity.* Journal of Clinical Investigation, 1998. **101**(11): p. 2351-2363.
59. Sukhova, G.K., et al., *Deficiency of cathepsin S reduces atherosclerosis in LDL receptor–deficient mice.* Journal of Clinical Investigation, 2003. **111**(6): p. 897-906.
60. Herroon, M.K., et al., *Macrophage cathepsin K promotes prostate tumor progression in bone.* Oncogene, 2012. **32**: p. 1580.
61. Wu, H., et al., *Cathepsin S activity controls injury-related vascular repair in mice via the TLR2-mediated p38MAPK and PI3K-Akt/p-HDAC6 signaling pathway.* Arteriosclerosis, thrombosis, and vascular biology, 2016.
62. Brömme, D., *Cathepsin V.* Handbook of Proteolytic Enzymes, 2013. **2**: p. 1831-1834.
63. Law, S., et al., *Identification of mouse cathepsin K structural elements that regulate the potency of odanacatib.* Biochemical Journal, 2016(2016).
64. Novinec, M., et al., *Interaction between human cathepsins K, L, and S and elastins: mechanism of elastinolysis and inhibition by macromolecular inhibitors.* Journal of Biological Chemistry, 2007. **282**(11): p. 7893-7902.
65. Novinec, M., et al., *Conformational flexibility and allosteric regulation of cathepsin K.* Biochemical Journal, 2010. **429**(2): p. 379-389.
66. Novinec, M., B. Lenarčič, and A. Baici, *Clusterin is a specific stabilizer and liberator of extracellular cathepsin K.* FEBS Letters, 2012. **586**(7): p. 1062-1066.
67. Nycander, M., et al., *Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop.* FEBS Letters, 1998. **422**(1): p. 61-64.
68. Voit, E.O., *A First Course In Systems Biology*. 1 ed. 2013, New York, NY: Garland Science, Taylor & Francis Group. 445-445.
69. Perlenfein, T.J. and R.M. Murphy, *A mechanistic model to predict effects of cathepsin B and cystatin C on β-Amyloid aggregation and degradation.* Journal of Biological Chemistry, 2017. **292**(51): p. 21071-21082.
70. Barry, Z.T. and M.O. Platt, *Cathepsin S cannibalism of cathepsin K as a mechanism to reduce type I collagen degradation.* The Journal of biological chemistry, 2012. **287**(33): p. 27723-30.
71. Novinec, M., et al., *A novel allosteric mechanism in the cysteine peptidase cathepsin K discovered by computational methods.* Nature communications, 2014. **5**: p. 3287-3287.

72. Sbongile, M. and M.E.S. Soliman, *In silico identification of irreversible cathepsin B inhibitors as anti- cancer agents: virtual screening, covalent docking analysis and molecular dynamics simulations.* Combinatorial chemistry & high throughput screening, 2015. **18**(4): p. 399-410.
73. Shah, P.P., et al., *Kinetic characterization and molecular docking of a novel, potent, and selective slow-binding inhibitor of human cathepsin L.* Mol Pharmacol, 2008. **74**(1): p. 34-41.
74. Thurmond, R.L., et al., *Identification of a potent and selective noncovalent cathepsin S inhibitor.* Journal of Pharmacology and Experimental Therapeutics, 2003. **308**(1): p. 268-276.
75. Tyagi, C., et al., *Mechanistic insights into mode of action of novel natural cathepsin L inhibitors.* BMC Genomics, 2013. **14**(Suppl 8): p. S10-S10.
76. Ma, S., L.S. Devi-Kesavan, and J. Gao, *Molecular dynamics simulations of the catalytic pathway of a cysteine protease: A combined QM/MM study of human cathepsin K.* J Am Chem Soc, 2007. **129**(44): p. 13633-13645.
77. Ferrall-Fairbanks, M.C., et al., *Computational predictions of cysteine cathepsin-mediated fibrinogen proteolysis.* Protein Science, 2017. **27**: p. 714-724.
78. Costa, M.G.S., et al., *How does heparin prevent the pH inactivation of cathepsin B? Allosteric mechanism elucidated by docking and molecular dynamics.* BMC Genomics, 2010. **11**(SUPPL. 5): p. 1-15.
79. Park, K.-Y., G. Li, and M.O. Platt, *Monocyte-derived macrophage assisted breast cancer cell invasion as a personalized, predictive metric to score metastatic risk.* Scientific reports, 2015. **5**: p. 13855-13855.
80. Hasegawa, C., et al., *Modeling and simulation of bone mineral density response from a phase 2 study of ONO-5334, a new cathepsin K inhibitor, to support dose selection in osteoporosis.* Journal of Clinical Pharmacology, 2014. **54**(8): p. 937-948.
81. Jakubzick, C.V., G.J. Randolph, and P.M. Henson, *Monocyte differentiation and antigen-presenting functions.* Nature Reviews Immunology, 2017. **17**: p. 349.
82. Bobryshev, Y.V., et al., *Macrophages and their role in atherosclerosis: Pathophysiology and transcriptome Analysis.* BioMed research international, 2016. **2016**: p. 9582430-9582430.
83. Park, K.-Y., W.A. Li, and M.O. Platt, *Patient specific proteolytic activity of monocyte-derived macrophages and osteoclasts predicted with temporal kinase activation states during differentiation.* Integrative Biology, 2012. **4**(12): p. 1459-1469.
84. Li, W.A., et al., *Detection of femtomole quantities of mature cathepsin K with zymography.* Analytical Biochemistry, 2010. **401**(1): p. 91-98.
85. Wilder, C.L., et al., *Manipulating substrate and pH in zymography protocols selectively distinguishes cathepsins K, L, S, and v activity in cells and tissues.* Archives of Biochemistry and Biophysics, 2011. **516**(1): p. 52-57.
86. Downey, G.P., et al., *Retention of leukocytes in capillaries: Role of cell size and deformability.* Journal of Applied Physiology, 1990. **69**(5): p. 1767-1778.
87. Steinman, R.M.B., Scott E.; Cohn, Zanvil A., *Membrane flow during pinocytosis. A stereologic analysis.* The Journal of Cell Biology, 1976. **68**(3): p. 665-687.

88. Beers, C., et al., *Differential regulation of cathepsin S and cathepsin L in interferon gamma-treated macrophages.* The Journal of experimental medicine, 2003. **197**(2): p. 169-79.
89. Gocheva, V., et al., *IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion.* Genes & Development, 2010. **24**(3): p. 241-255.
90. Palmieri, M., et al., *Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways.* Human Molecular Genetics, 2011. **20**(19): p. 3852-3866.
91. Turk, B., D. Turk, and V. Turk, *Lysosomal cysteine proteases: More than scavengers.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 2000. **1477**(1): p. 98-111.
92. Qi, X., et al., *Cathepsin B modulates lysosomal biogenesis and host defense against Francisella novicida infection.* The Journal of Experimental Medicine, 2016: p. 1-17.
93. Matsumoto, K., et al., *Structural basis of inhibition of cysteine proteases by E-64 and its derivatives.* Biopolymers - Peptide Science Section, 1999. **51**(1): p. 99-107.
94. Menard, R., et al., *Autocatalytic processing of recombinant human procathepsin L.* Journal of Biological Chemistry, 1998. **273**(8): p. 4478-4484.
95. Riese, R.J., et al., *Cathepsin S activity regulates antigen presentation and immunity.* The Journal of clinical investigation, 1998. **101**(11): p. 2351-2363.
96. Muntener, K., et al., *Exon skipping of cathepsin B: mitochondrial targeting of a lysosomal peptidase provokes cell death.* J Biol Chem, 2004. **279**(39): p. 41012-7.
97. Tamhane, T., et al., *Nuclear cathepsin L activity is required for cell cycle progression of colorectal carcinoma cells.* Biochimie, 2016. **122**: p. 208-18.
98. Turk, V., et al., *Cysteine cathepsins: From structure, function and regulation to new frontiers.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2012. **1824**(1): p. 68-88.
99. Brömme, D., *Cathepsin K.* Handbook of Proteolytic Enzymes, 2013. **2**: p. 1801-1807.
100. Brubaker, K.D., et al., *Cathepsin K mRNA and protein expression in prostate cancer progression.* Journal of Bone and Mineral Research, 2003. **18**(2): p. 222-230.
101. Reddy, V.Y., Q.Y. Zhang, and S.J. Weiss, *Pericellular mobilization of the tissue-destructive cysteine proteinases, cathepsins B, L, and S, by human monocyte-derived macrophages.* Proceedings of the National Academy of Sciences, 1995. **92**(9): p. 3849.
102. Beers, C., et al., *Differential regulation of cathepsin S and cathepsin L in interferon γ–treated cacrophages.* The Journal of Experimental Medicine, 2003. **197**(2): p. 169-179.
103. Ferrall-Fairbanks, M.C., *Effects of cathepsin proteolytic network dynamics on extracellular matrix degradation in biological machines and invasive disease*, in *Biomedical Engineering*. 2017, Georgia Institute of Technology: Atlanta, Georgia.
104. Lake-Bakaar, G., et al., *Measurement of trypsin in duodenal juice by radioimmunoassay.* Gut, 1980. **21**(5): p. 402.

105. Aoshima, H., et al., *Analysis of autodegradation sites of thermolysin and enhancement of its thermostability by modifying Leu155 at an autodegradation site.* The Journal of Biochemistry, 2004. **135**(4): p. 547-553.
106. Hoops, S., et al., *COPASI - A COmplex PAthway SImulator.* Bioinformatics, 2006. **22**(24): p. 3067-3074.
107. Littlewood-Evans, A.J., et al., *The Osteoclast-associated protease cathepsin K is expressed in human breast carcinoma.* Cancer Research, 1997. **57**(23): p. 5386.
108. Ferrall-Fairbanks, M.C., et al., *PACMANS: A bioinformatically informed algorithm to predict, design, and disrupt protease-on-protease hydrolysis.* Protein Science, 2017. **26**(4): p. 880-890.
109. Akatov, V.S., et al., *Low pH value of pericellular medium as a factor limiting cell proliferation in dense cultures.* Experimental Cell Research, 1985. **160**(2): p. 412-418.
110. Daigneault, M., et al., *The identification of markers of macrophage differentiation in PMA-stimulated THP-1 cells and monocyte-derived macrophages.* PLoS ONE, 2010. **5**(1).
111. Tiwari, P.M., et al., *Engineered mRNA-expressed antibodies prevent respiratory syncytial virus infection.* Nat Commun, 2018. **9**(1): p. 3999.
112. Porath, J., et al., *Metal chelate affinity chromatography, a new approach to protein fractionation.* Nature, 1975. **258**.
113. Yan, D., et al., *STAT3 and STAT6 Signaling Pathways Synergize to Promote Cathepsin Secretion from Macrophages via IRE1alpha Activation.* Cell Rep, 2016. **16**(11): p. 2914-2927.
114. Ménard, R., et al., *Autocatalytic processing of recombinant human procathepsin L: contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro.* Journal of Biological Chemistry, 1998. **273**(8): p. 4478-4484.
115. Yan, D., et al., *STAT3 and STAT6 signaling pathways synergize to promote cathepsin secretion from macrophages via IRE1α activation.* Cell Reports, 2016. **16**(11): p. 2914-2927.
116. Howlader, N., et al., *SEER cancer statistics review, 1975-2014*. 2017: Bethesda, MD.
117. Black, W.C. and M.D. Percival, *The consequences of lysosomotropism on the design of selective cathepsin K inhibitors.* ChemBioChem, 2006. **7**(10): p. 1525-1535.
118. Falgueyret, J.P., et al., *Lysosomotropism of basic cathepsin K inhibitors contributes to increased cellular potencies against off-target cathepsins and reduced functional selectivity.* Journal of Medicinal Chemistry, 2005. **48**(24): p. 7535-7543.
119. Rünger, T.M., et al., *Morphea-like skin reactions in patients treated with the cathepsin K inhibitor balicatib.* Journal of the American Academy of Dermatology, 2012. **66**(3).
120. Soetaert, K., T. Petzoldt, and R.W. Setzer, *Solving differential equations in R: Package deSolve.* Journal of Statistical Software, 2010. **33**(9): p. 1-25.
121. Lee, A.V., S. Oesterreich, and N.E. Davidson, *MCF-7 cells—Changing the course of breast cancer research and care for 45 years.* Journal of the National Cancer Institute, 2015. **107**(7): p. 1-4.

122. Qu, Y., et al., *Evaluation of MCF10A as a reliable model for normal human mammary epithelial cells.* PLoS ONE, 2015. **10**(7): p. 1-16.
123. Kathryn, J.C., G. Sireesha V, and L. Stanley, *Triple negative breast cancer cell lines: One tool in the search for better treatment of triple negative breast cancer.* Breast Dis, 2012. **32**: p. 35-48.
124. Nomura, T. and N. Katunuma, *Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells.* The Journal of Medical Investigation, 2005. **52**(1,2): p. 1-9.
125. Kieslich, C.A., F. Boukouvala, and C.A. Floudas, *Optimization of black-box problems using Smolyak grids and polynomial approximations.* Journal of Global Optimization, 2018. **71**(4): p. 845-869.
126. Desmarais, S., et al., *Effect of cathepsin k inhibitor basicity on in vivo off-target activities*. Vol. 73. 2008. 147-56.
127. Peroni, A., et al., *Drug-induced morphea: Report of a case induced by balicatib and review of the literature.* Journal of the American Academy of Dermatology, 2008. **59**(1): p. 125-129.
128. Kirkegaard, T. and M. Jäättelä, *Lysosomal involvement in cell death and cancer.* Biochimica et Biophysica Acta - Molecular Cell Research, 2009. **1793**(4): p. 746-754.
129. Pareja, F., et al., *Breast cancer heterogeneity: Roles in tumorigenesis and therapeutic implications.* Current Breast Cancer Reports, 2017. **9**(1): p. 34-44.
130. Olson, O.C., et al., *Tumor-associated macrophages suppress the cytotoxic activity of antimitotic agents.* Cell Reports, 2017. **19**(1): p. 101-113.
131. Roos, A. and W.F. Boron, *Intracellular pH.* Physiological Reviews, 1981. **61**(2).
132. Barrett, A.J., et al., *L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L.* Biochem. J, 1982. **201**: p. 189-198.
133. Montenez, J.P., et al., *Increased activities of cathepsin B and other lysosomal hydrolases in fibroblasts and bone tissue cultured in the presence of cysteine proteinases inhibitors.* Life Sciences, 1994. **55**(15): p. 1199-1208.
134. Turk, B., et al., *Acidic pH as a physiological regulator of human cathepsin L activity.* European Journal of Biochemistry, 1999. **259**(3): p. 926-932.
135. Yokota, S. and K. Kato, *The heterogeneity of rat kidney lysosomes revealed by immunoelectron microscopic staining for cathepsins B and H.* Histochemistry, 1988. **89**(5): p. 499-504.
136. Pham, P.L., et al., *Large-scale transient transfection of serum-free suspension-growing HEK293 EBNA1 cells: Peptone additives improve cell growth and transfection efficiency.* Biotechnology and Bioengineering, 2003. **84**(3): p. 332-342.
137. Cirman, T., et al., *Selective disruption of lysosomes in HeLa cells triggers apoptosis mediated by cleavage of Bid by multiple papain-like lysosomal cathepsins.* Journal of Biological Chemistry, 2004. **279**(5): p. 3578-3587.
138. Droga-Mazovec, G., et al., *Cysteine cathepsins trigger caspase-dependent cell death through cleavage of bid and antiapoptotic Bcl-2 homologues.* Journal of Biological Chemistry, 2008. **283**(27): p. 19140-19150.

139. Quintanilla-Dieck, M.J., et al., *Expression and regulation of cathepsin K in skin fibroblasts.* Experimental Dermatology, 2009. **18**(7): p. 596-602.
140. Luo, H., et al., *Cancer-associated fibroblasts: A multifaceted driver of breast cancer progression.* Cancer Letters, 2015. **361**(2): p. 155-163.
141. Srisa-Art, M., et al., *Monitoring of real-time streptavidin−biotin binding kinetics using droplet microfluidics.* Analytical Chemistry, 2008. **80**(18): p. 7063-7067.
142. Milo, R. and R. Phillips, *Cell Biology by the Numbers*. 2016, New York, NY: Garland Science.
143. Wegmeyer, H., et al., *Mesenchymal stromal cell characteristics vary depending on their origin.* Stem Cells Dev, 2013. **22**(19): p. 2606-18.
144. Wang, L.T., et al., *Human mesenchymal stem cells (MSCs) for treatment towards immune- and inflammation-mediated diseases: review of current clinical trials.* J Biomed Sci, 2016. **23**(1): p. 76.
145. Zhukareva, V., et al., *Secretion profile of human bone marrow stromal cells: donor variability and response to inflammatory stimuli.* Cytokine, 2010. **50**(3): p. 317-21.
146. Gotherstrom, C., et al., *Difference in gene expression between human fetal liver and adult bone marrow mesenchymal stem cells.* Haematologica, 2005. **90**: p. 1017-1026.
147. Kim, J.M., et al., *Comparative secretome analysis of human bone marrow-derived mesenchymal stem cells during osteogenesis.* J Cell Physiol, 2013. **228**(1): p. 216-24.
148. Sheen-Chen, S.-M., et al., *Serum concentration of tumor necrosis factor in patients with breast cancer.* Breast Cancer Research and Treatment, 1997. **43**(3): p. 211-215.
149. Hussein, M.Z., et al., *Serum IL-6 and IL-12 levels in breast cancer patients.* (1110-4902 (Print)).